

HELSINGIN YLIOPISTO

# Use of the Latin script in non-Latin script languages

---

Anni Luoma  
Master's Thesis  
Linguistic Diversity in the Digital Age  
Faculty of Arts  
University of Helsinki  
5/2019



Tiedekunta/Osasto – Fakultet/Sektion – Faculty Faculty of Arts		Laitos - Institution - Department Department of Languages	
Tekijä - Författare - Author Anni Luoma			
Työn nimi - Arbetets titel - Title Use of the Latin script in non-Latin script languages			
Oppiaine - Läroämne - Subject General Linguistics			
Työn laji - Arbetets art - Level Master's Thesis		Aika - Datum - Month and year May 2019	Sivumäärä - Sidoantal – Number of pages 72 + 18
Tiivistelmä - Referat - Abstract <p>This study examines the use of the Latin script in languages where the Latin script is not the official writing system. The aim of the research is to get a general perception of where and why the Latin script is used in those languages. The study concentrates on the use of the Latin script of individual writers, especially in informal writing situations online. The study also examines if the individuals have seen the Latin script used by others. In addition, the study finds out about the keyboard functionality when writing in the Latin and non-Latin scripts.</p> <p>Many case studies have covered the phenomenon of digraphia, which is a situation where one language uses two writing systems. However, a broader crosslinguistic study has not been done on this topic. The Latin script is the dominant script on the Internet which might have an impact on the script choice in online writing. Internet language and chatting are broadly studied but because of the continuously changing nature of the Internet, more research is required on these topics. This study aims to fill in some of the gaps that are left open by previous research.</p> <p>I gathered data by an online questionnaire and by four semi-structured interviews. By means of the questionnaire, I got open question answers and quantitative data from 142 respondents. Additionally, I had four interviews with representative users of different writing systems to get more detailed information about the use of the Latin script. The data was analysed and compared with previous research from the point of view of five groups of different scripts.</p> <p>The general finding of this study is that the respondents choose to write in the Latin script mostly when the non-Latin keyboard is not available or when fast and easy informal writing style is required. The Latin script has many different functions. It is mostly used in online chatting, texting with friends or when one's personal information needs to be written in travelling documents. However, people prefer to write in the official non-Latin script even though it might sometimes feel difficult or slow to use. The study points out that it cannot be predicted whether a respondent uses the Latin script or not, since the results are mostly very variable. The need or even the eagerness to use the Latin script when it is not the official script, reflects the functions and facilities that are available or unavailable for different scripts.</p>			
Avainsanat – Nyckelord - Keywords Latin script, writing systems, digraphia, script-switching, linguistic landscapes, Internet writing, sociolinguistics			
Säilytyspaikka - Förvaringsställe - Where deposited Helsinki University Main Library			
Muuta tietoa - Övriga uppgifter - Additional information			

# Table of Contents

<b>1 Introduction</b>	1
<b>2 Theoretical background</b>	5
2.1. Writing systems	5
2.2. Internet writing	7
2.2.1. ASCII, ISO and Unicode	8
2.2.2. Texting and binary codes	8
2.2.3. Romanisation	9
2.2.4. Keyboards	10
2.2.5. Code-switching vs. script-switching	11
2.3. Choosing the script and writing style	13
2.3.1. Social reasons	13
2.3.2. Technical reasons	14
2.3.3. Individual reasons	14
2.3.4. Language political reasons	15
2.4. Societal aspects of script choice	15
2.4.1 Digraphia vs. diglossia	16
2.4.2. Scripts and language politics	17
2.4.3. Linguistic landscape	17
2.5. On the hypotheses	18
<b>3 Methods and Data</b>	19
3.1. The Online questionnaire	19
3.2. The Interviews	21
3.3. The Data	23
3.4. Research ethics	26
<b>4 Results</b>	28
4.1. Online questionnaire results	28
4.1.1. Use of the Latin script by the respondent	28
4.1.2. Use of the Latin script by others	35
4.1.3. What is written in the Latin script?	37
4.1.4. Writing on electronic devices	38
4.2. The Interviews	41
4.2.1. Russian (Cyrillic)	41
4.2.2. Farsi (Arabic)	42
4.2.3. Cantonese (traditional Chinese characters)	43
4.2.4. Thai (Thai)	45
4.3. To conclude on results	46

<b>5 Analysis .....</b>	<b>47</b>
5.1. Cyrillic and Latin scripts .....	47
5.1.1. Russian, Bulgarian and Chechen.....	47
5.1.2. Serbia with digraphia .....	48
5.2. Arabic and Latin scripts .....	51
5.3. Chinese and Latin scripts .....	54
5.3.1. Simplified characters (Mainland China) .....	54
5.3.2. Traditional characters (Taiwan) .....	55
5.3.3. Traditional characters (Hong Kong) .....	56
5.4. Thai and Latin scripts .....	57
5.5. Others and Latin scripts.....	58
5.5.1 Japanese.....	59
5.5.2. Hindi.....	60
5.5.3. Miscellaneous.....	60
5.6. Comparing and contrasting the scripts .....	61
<b>6 Discussion.....</b>	<b>65</b>
6.1. On the methods.....	65
6.2. On the results.....	67
<b>7 Conclusions .....</b>	<b>69</b>
References .....	73
Attachment 1 .....	77
Attachment 2 .....	83

# 1 Introduction

This research is about the habits of the choice of writing system in languages, where at least one of the official writing systems is a non-Latin-alphabet based writing system. The aim of the research is to get a general perception of the use of the Latin or Roman script in languages where it is not the conventional choice of writing system. The main research questions of this research are: 1) Do the writers use the Latin script if their language's official writing system is a non-Latin script? 2) If yes, when and why do they do that?

The study investigates three research hypotheses. The first hypothesis is that all the respondents either sometimes write their language in the Latin alphabet or they have at least seen it written in the Latin alphabet somewhere by others. The second hypothesis is that the main domains of use of the Latin script are when writing something online or with an electronic device. The third hypothesis is that in some cases writing with the Latin script is voluntary or even desirable and, in some cases the Latin script is used only because the keyboard for the official writing system is not available.

These hypotheses are motivated by the phenomenon of globalism and the continuously changing interlingual Internet. One might assume that the people who use the Internet, have not avoided the dominance of English language online and therefore they have been at least in some contact with the Latin alphabet. Chatting is often informal, and it is used for fast communication between friends and family. Therefore, an unconventional script choice might be a suitable for informal communication. Additionally, the laziness to switch keyboards can be a reason for writing in the Latin script. More about the motives for the hypotheses is discussed in the Chapter 2.5.

In this study I concentrate on the Latin script use in situations where the respondent uses both the Latin script and a non-Latin script or just the Latin script. Therefore, I take into consideration all kinds of written language use as long as some part of the target language (officially written in the non-Latin script) is written in the Latin script. The research takes into consideration single words written in the Latin alphabet, as well as complete texts and everything in between these two. The research data consists of online questionnaire responses of 142 respondents, who speak languages that officially use a non-Latin script, and interviews of four of the respondents of the questionnaire. Therefore, the data is not actual texts written in the Latin or non-Latin scripts. The data consists of answers given by the respondents on

questions about their Latin script use and other related matters. The research considers mostly the language use of an individual on the Internet, since there the language use is mostly free of strict rules of grammar or orthography. In addition, the Latin script use by groups of people, administration etc. is discussed. Therefore, this study does not examine only the Latin script use of an individual but it also aims to find out how the individual has seen the script used by others.

The research compares the use of different writing systems. The collected data is divided into five groups: “Cyrillic”, “Arabic”, “Chinese”, “Thai” and “Others”. The division of the first four groups are inspired by the writing systems that are used in the languages of the four interviewees. The group “Others” consists of all the other writing systems. In the following chapters these five groups are examined individually and compared with each other. One of the aims of this research is to see if the official writing system of a language has any effect on the use of the Latin script.

There is no existing research that would look into these questions from a broad crosslinguistic perspective. This research aims to broaden the information about the phenomenon of switching between writing systems when writing a language even though the study does not cover every writing system that exists. This research gives an up-to-date view of the script-switching habits around the globe. For this study, 142 respondents provided the data that is used for making generalisations. However, internet mediated language changes constantly and in ten years from now, the results might be completely different than what they are now.

The written language as a field of sociolinguistic studies is a rather young phenomenon and the written language on the Internet an even younger still. There is a lot of research on code-switching and language choice on the Internet but less research about the script choice online. Scholars like Ivković (2015), Lee (2007), Danet & Herring (2007), Tseliga (2007), Koutsogiannis (2007) and Androutsopoulos (2012) have done studies especially on the Latin script use online in languages that officially utilise a non-Latin script. I use these studies as background literature for this study and then compare the results with the data that was gathered for this study.

There are several sources of motivation for this study. As mentioned previously, there are only case studies or theoretical studies done on this matter. Crosslinguistic and global studies should be done in order to see the phenomenon of script-switching in a larger scale. Internet

language is continuously shifting its form and it continues to sprout new varieties of ways to communicate (Herring, 2007). The possible research topics concerning internet mediated language are basically countless and becoming more uncountable while the electronic communication habits continue to evolve. Internet has become a great platform for all kinds of written language exchange, which makes it an interesting and always current topic of research. The Internet is closely connected to smartphones, computers, tablets and other electronic devices that provide the means to write and send messages. The technologies that enable us to use creative ways to communicate online should be studied more. Different keyboard layouts, input systems and text editors are constantly developing, and it should be ensured that all language scripts are taken into account while doing so. Script choice related topics, such as nationalism, language identity and using language politics as a means to rule language communities are interesting and relevant topics for study as well.

Writing in general is mainly done out of social motivations. When another person is not in the range of hearing, messaging is often the solution for getting the information to its destination. Writing a letter is nowadays considered a personal and thoughtful act that demands some time and energy from the sender. Texts, emails and other phone or computer mediated formats on the other hand are fast and reliable means to get the message sent to the recipient. In all the ways of writing to another person, the writer's aim is usually to get the message across and to be understood. Therefore, some thought considering the recipient is required. The writer must know in which language, in which style or in which script s/he needs to write so that the reader understands the message in the originally intended way. There are several ways to make this written communication more interesting or socially bonding. Inside jokes, code-switching as well as script-switching add some special personal touch into the text for the reader to decode and enjoy.

There is large number of sociolinguistic factors that can be discovered from an individual's choice of a script. Therefore, it is relevant to investigate these connections between an individual and the surrounding sociolinguistic factors. Choosing a certain script can tell about the need to belong in a group. The script choice can also tell that the writer wants to take an acquaintance living abroad into consideration. The Latin script can work as a means for understanding or on the contrary, work as a secret code. This study gives some answers to these emerged deliberations and demands that need constant response. One of the things that the study does not concentrate on is the phonology on languages and their effect on the

suitability of writing systems to each language even though it might be a relevant topic for study.

After this introduction chapter, I present some technical terms and background knowledge in Chapter 2. There I go through factors that affect the choice of script from different points of views. I also explain briefly some phenomena related to scripts and script research. In Chapter 3 I introduce the methods used in this study as well as present the general information of the data. The research ethical factors are discussed right after the data introduction. In Chapter 4 I present the most relevant results that came up from the data. The detailed quantitative data is presented with the aid of figures and tables. The open question answers specify the results of the quantitative data. The data of the five groups of different scripts is discussed. The interview data is presented as compatible texts with a narrator perspective. Chapter 5 concentrates on analysing the quantitative data of the online questionnaire and the qualitative data which consists of the interviews and the open question answers gotten from the questionnaire. I introduce the tendencies and similarities as well as point out some differences that emerged in the results. In addition, in the analysing chapter the groups of different scripts are compared. Chapter 6 is where I discuss the suitability of the methods for this study and consider the results of the study in comparison to the previous literature. In the final Chapter 7 I conclude the main findings of this research and go through the research questions and their newly discovered answers. I return to the hypotheses and report if they are supported by the data.



## 2 Theoretical background

In this chapter, some of the basic terminology and concepts related to this study are explained. Throughout this study script, alphabet and writing system are used carrying the same meaning. In this study I use the term Latin script instead of the Roman script entirely out of personal preference. However, the term latinisation is not used because it often has the meaning of translating names into Latin language, which is not a topic related to this study. Instead, romanisation is discussed in Subchapter 2.2.3.

### 2.1 Writing systems

There are hundreds of writing systems in the world, which can be crosslinguistically divided into logographic and phonographic systems. The logographic systems include morphemic systems and the polymorphemic based systems. In these systems the units in the writing system can represent single morphemes or even complete words. The phonographic divides into syllabic, segmental and featural systems. The phonographic writing systems include graphemes that represent phonological units of the language. (Sampson 2015: 24–25)

The Latin alphabet is often a popular choice for a languages new writing system because it is easy to write, it is phonemic and the words written in the Latin script are simple to read (Grivelet 2001a: 1, Hosken 2003: 6). On the other hand, some scripts are in danger of falling out of use because the dominant language of the country utilises the Latin script or because the non-Latin writing systems are not properly taught at school (Anderson 2005: 28).

The written symbols can represent the sounds in various ways depending on the language where the script is applied. In English, which utilises the Latin alphabet, the letter <c> represents phonemes /k/ (as in “cat”) or as /s/ (as in “cent”). In Castilian Spanish, however <c> represents the phoneme /θ/ and in Turkish it is marks the phoneme /dʒ/. The languages that use the Latin alphabet have around 26–29 letters that are used as language specific phonetic representations. (Palfreyman & Al Khalil 2007: 44–45)

In many languages, there are standards on how words should be written, when it comes to official texts as there are orthographical rules that must be followed. In official texts misspelling of words can lead to misunderstandings. The common spelling rules give words standard shapes that are easy to read and understand. Homophones exist but depending on the

language and its orthography they might be written in different ways. Homonyms, words that have the same shape, can be memorised and kept in mind to avoid ambiguity.

The writing systems can be described to have shallow or deep orthography or something from between. In shallow orthographies the spelling corresponds with the sounds of the spoken languages. In deep orthographies, on the opposite, the pronunciation is not as clearly visible as in shallow orthographies. (Ellis 2004: 438)

Ideally, a language's orthography is fitting for the language in a way that it is easy to use by the readers and writers. However, language and its orthography are not the same thing. Natural languages are structures that have evolved during thousands of years, merged, diverged and assumed new functions and taken over new domains. The orthography of a language is a tool for writing the natural language. It is artificial and the rules are conventions that have been designed and agreed on before taking them into broader use.

Orthography is likewise a different thing than script, since several languages can have the same script but then use different orthographical rules to write with it (Hansell 2015). Script or writing system consists of alphabet, characters or other figures that decode the spoken language into a written form. In addition, the script only aims to represent the language, not to be the language (Sampson 2015: 18). Therefore language, orthography and script are all separate systems with their own functions and natures.

The language specific orthographic rules do not always guide the person when writing online. Writing in various Internet settings encourages the writers to write more phonemically transparent text in addition to writing more economically, ignoring some of the "unnecessary" orthographic features on purpose (Silva 2013: 155). After continuous use of the Latin script by online users, even some orthographic conventions can be established (Ivković 2013: 336). Additionally, in today's Internet the writing speed is key: grammatical correctness or writing complete words is secondary to the need to be efficient in speed (Silva 2013: 156). In online setting the writer has more freedom when it comes to orthographical rules.

Now I shortly introduce some writing systems that are relevant from the point of view of this study. The interviewees represent speakers of four languages: Russian, Farsi, Cantonese Chinese and Thai. Russian is written with the Cyrillic alphabet which is a relative writing system to the Latin alphabet. They are phonographic writing systems where a phoneme is

preferably marked with one grapheme. Because of the similar nature of these writing systems, they are easily interchangeable in many cases. The modern Farsi utilises a modified form of the Arabic writing system, which is an "impure" Abjad that marks mostly only consonants but also has convenient ways to mark the vowels if needed. Cantonese Chinese usually uses the traditional Chinese characters, which is a logographic system and the parent writing system to simplified Chinese characters that are in use in mainland China. Thai language uses an Abugida script in which consonant–vowel combinations are written with one letter. In addition to these four writing systems, later in the text also the writing systems used by Japanese and Hindi speakers are discussed more. Japanese utilises a combination of three scripts: syllabic Hiragana and Katakana with addition of logographic Kanji. The Kanji script is adopted from the Chinese Hanzi characters. Hindi is written with the Devanagari script. It is an Abugida script that composes a letter by combining a consonant with a vowel. In this study, other scripts are discussed as well, but their characteristics are discussed later when needed.

## **2.2 Internet writing**

Computer-mediated communication, which consists for example of e-mails, forums, blogs and chats, happens within specific rules for production and reception (Silva 2013: 150). On the other hand, the Internet is not as strict of a language domain as some of the “real life” writing settings can be. Chat users make mistakes in their writing at such regularity that the online writing might actually be an intention to create the ideal writing system (Silva 2013: 147–148). When a person writes to a chat group, s/he can be quite sure that her/his grammar teacher is not there to be marking the chat text with a red pen. Therefore, switching the writing system mid-sentence would be more acceptable in certain Internet platforms compared to for example writing school essays.

Internet has the possibility to either save or repress the diversity of different scripts (Anderson 2005: 27–28). The dominant writing systems such as the Latin and Cyrillic scripts can override less used scripts if they are not sufficiently supported by electronic devices. On the other hand, if they are, Internet is a place where language users can increase their knowledge on the script use of their language.

### **2.2.1 ASCII, ISO and Unicode**

The ASCII (American Standard Code for Information Interchange) was originally invented for the basic computational interaction and it was created from the basis of the English alphabet (ASCII 2019). The ASCII can read the letters A–Z in upper and lower case, digits from 0–9, some punctuations, a few special characters and 33 values like space and line feed (Zentgraf 2015). The conciseness of the ASCII system has affected the accessibility of the Internet. Therefore, the world of electronic devices has given English a prestige status online compared to the languages that use other characters than those that are provided by ASCII (Danet & Herring 2007: 9). English being the dominant language online profits the people who use English as their mother tongue but makes expressing more difficult to those who have some other mother tongue (Axelsson & Abelin 2007: 363).

Shortly after the ASCII became the standard in internet communication, the International Organization for Standardization (ISO) started to adapt the ASCII code also for other languages that utilise letters like ö, â and ü. First, some ISO extensions to the Latin letters were added, later followed by extensions for writing letters in Hebrew, Arabic and Russian etc. Then the 32-bit code called the Universal Coded Character Set (UCS) was launched. It was supposed to become the dominant encoding scheme for international computing, but Unicode (the Unification Code) overruled it with its simplicity. (Maxfield 2011)

Unicode was founded because of the need to decode any character on any platform and for any language instead of using different schemes for each language. Since the first version of the Unicode Standard in 1991, the Unicode had enabled computational writing for over 150 writing systems by the end of 2018 (Unicode Standard 2018). The encoding of Unicode has been standardised internationally and it is similar in every country which results to the fact that computing in various languages and scripts has become less and less problematic for the non-Latin script users all over the world (Anderson 2005: 27).

### **2.2.2 Texting and binary codes**

The GSM 03.38 is a character set that extends the ASCII characters that are used in many European languages (Connection Software 2018). Primarily, the standard SMS messages use the GSM 03.38 characters, but for languages like Chinese and Arabic the SMS message is

encoded in the Unicode because these languages use characters that are not in the GSM 03.38 character set (World Text n.d.).

Computers handle written characters in binary codes that consist of bits 1 and 0. These two bits are used in a bit string of seven (for ASCII and GSM 03.38) and eight (for Unicode). These strings represent the characters that are written on the computer. In the ASCII encoding one bit string encodes one character but for example in Unicode, one character might require up to four bit strings. (Zentgraf 2015)

If the SMS message is written in a script, that needs to be encoded with Unicode, the messages can contain up to 70 characters. In comparison to for example Finnish, which can be encoded with the less space consuming GSM 03.38 character set, one message can contain up to 160 characters. This is because of the Unicode binary codes are longer than the ones of the GSM 03.38. Writing with the Latin alphabet and not with other scripts therefore prevents from paying for additional messages since the Latin script allows the writer to include more characters in one text message. (World Text n.d.)

### **2.2.3 Romanisation**

Romanisation means the substitution of non-Latin graphemes with the corresponding Latin alphabet (Ivković 2015). For the benefit of the English-speaking world, romanisation conventions have been made to dozens of languages, and they are adopted and made official usually by the UN or the BGN (the United States Board on Geographic Names) and PCGN (Permanent Committee on Geographical Names for British Official Use) (UNGEGN 2014). In addition, ISO has made its own romanisation standards for many languages (ISO n.d.). In some languages, the romanised alphabets are mostly for transliterating the language for international use rather than for example for chatting (Ivković 2015). Therefore, in some languages the use of the Latin script in online communication is a secondary domain of use that could have started, in some cases, from the need of regular people to communicate faster and easier on the Internet.

Romanisation has often been perceived as a negative thing that endangers local cultural identities (Spasov 2012: 1487). The dominant status of the Latin script can marginalise smaller scripts. The Latin script can be perceived as a sign of western influence, which is not always welcome.

## 2.2.4 Keyboards

There are several different Latin alphabet-based keyboards available for electronic devices. The most common ones are QWERTY, QWERTZ, AZERTY, Colemak and Dvorak. In addition to these, keyboards with language specific diacritics and letters are available. Smaller non-Latin scripts are not as widely used on the Internet as the Latin script. Therefore, for many small non-Latin writing systems there are inevitably fewer options when it comes to choosing a keyboard.

On smartphones it is very common to switch between different keyboards on different scripts. Arguably every smartphone user has at least switched the keyboard into emoji mode. On computers with a physical keyboard the switching between keyboards is more arduous since the letter stickers do not change on the keys by pushing a button.

People living in diaspora need to adjust their writing to the technological options available in the country of residence. This might mean the necessity to use the keyboard that is available and not the keyboard that would be needed. The writers in such situation can also bring some new additional orthographical writing solutions to the linguistic community. Some minority ethno-linguistic groups might even start practices in writing that are used only within the community in question. (Ivković 2015)

In Figure 1 there are keyboards that the interviewees of this research use on their phones (which are not all iPhones). The first one is the Russian language keyboard with upper case letters. The second is the Arabic script keyboard that is also used by Farsi speakers. The third is the Cangjie keyboard that can be used for writing Cantonese Chinese. The fourth keyboard is the Thai language keyboard.



Figure 1. Four different keyboards on iPhone 5s.

The logic of the third keyboard Canjie is based on radicals, a.k.a. smaller components of a character. The Canjie keyboard predicts the possible options for characters on the basis of the order of pressing the keys that represent the radicals. Cantonese can also be written on a QWERTY based keyboard or on a keyboard that accepts finger gestures. As mentioned by some of the respondents, people choose the keyboard according to their availability on their phone, their knowledge of typing and the type of the text (more of these results are seen later in the text). Some keyboards cannot fit all of the alphabets or characters on one “page” of a keyboard. In order to type some scripts, one needs to press multiple keys to get the preferred grapheme. In the English Latin keyboard or the Russian Cyrillic keyboard, however, the main letters can fit on one page and often only one press of a key is enough to create the right grapheme.

Some languages employ a Latin script-based typing system on electronic devices for writing a non-Latin script. For example the Romaji system is widely used for writing Japanese and Pinyin for writing Mandarin Chinese in mainland China. In these cases, the Latin script works as an auxiliary script that is based on the phonetic representation of the language. Even though the words are typed in the Latin script, the outcome on the screen is in another script, for example in simplified Chinese characters. In Mandarin Chinese it is rather easy to type words in the Pinyin system (for example “ni hao”, typed without tone marking) and then choose on the screen the wanted characters that correspond with the Pinyin representation (in this case “你好”). The outcome on the screen therefore is in the Chinese characters. As seen from the results of this research later in the text (Chapter 4), the Latin script is used by many of my respondents especially when using electronic devices.

### **2.2.5 Code-switching vs. script-switching**

Script-switching is an instance where two scripts are used simultaneously in the same text with one or several words written in one script and the other words written in another script. The words may or may not consist of code-switched or borrowed words. (Panović 2018: 73)

Usanova (2019) introduces two types of scriptural mixing: interscriptual mixing and intrascriptual mixing. The former type, interscriptual mixing, is when a person combines graphemes from two different scripts for example the Latin script and the Cyrillic script. The latter one, intrascriptual mixing, is when a person substitutes some graphemes with other graphemes from the same script. (Usanova 2019: 123)

The motivation behind socio-linguistically related choices in multilingual settings can be the same for both speech and writing, which makes code-switching and script-switching similar kind of actions (Angermeyer 2005: 515). Code-switching usually occurs only in informal settings (Angermeyer 2005: 494) which means that the Internet communication would be a potential domain for code-switching and therefore also to script-switching.

Script-switching is not always informal since some abbreviations or names of businesses are often written in their original form. Words like iPhone, AIDS, Finnair and WTC are sometimes not translated or transcribed to a non-Latin script form. If a non-Latin script language officially uses the Latin script forms of this kind of words, script-switching can be used in formal settings too.

Code-switching occurs for example in young people's everyday conversation, where a second language is added in various places to the flow of the speech. Code-switching in a conversation can occur for example in Finnish and English as in the example sentence 1:

- (1) "Meillä oli eilen *meeting*, se oli aika *perfect*." <sup>1</sup>  
 "We had a meeting yesterday, it was pretty perfect."

Script-switching is a similar kind act where two different "codes" are used in the conversation. The difference is that code-switching occurs in spoken language and script-switching occurs in written language. In the example sentence 2, script-switching happens between the Cyrillic and the Latin scripts. The word written in the Latin script is a name of a foreign company that has not been translated to Russian nor transcribed to the Cyrillic letters. The third example sentence is also script switching, but the whole sentence is in Cantonese and does not include foreign words.

- (2) "Завтра мы полетим на самолете *Ryanair*."  
 "Tomorrow we fly on Ryanair plane."  
 (3) "貓貓係到 *lam lam* 𦏧。"  
 "The cat is licking." (Free translation.)

The motivations to use code-switching or script-switching are sometimes possibly the same. They are used for example for fun, to appear international or because the word came to mind faster in the unconventional language or script (as in example 3). Code-switching and script-

---

<sup>1</sup> This example and the following two examples are fabricated sentences. The first one is invented by me, the second by the Russian speaking interviewee and the third by the Cantonese speaking interviewee.



switching can occur in many other forms than in those presented in the three examples, but here only three examples are presented in order to give a general perception of the phenomenon.

Since this research does not utilise naturally produced texts as data, the script-switching investigation is very limited in this study. However, some script-switching related questions were included in the online questionnaire. Script-switching is therefore included in this study but from the perspective of the respondents.

## **2.3 Choosing the script and writing style**

When writing in any language, a script must be chosen by the writer. Several things affect this choice and they can come from the surrounding circumstances or from personal motives.

### **2.3.1 Social reasons**

The multilingual internet constantly triggers the writer to choose a language and the choice is dependent on the technological, sociocultural and political settings of the given situation (Danet & Herring 2007: 17). Considering all the pros and cons of a certain writing system, the writer is unconsciously influenced in the choice of the writing system (Silva 2013: 149). The language is always selected and fitted to the purpose of its use as there is no moderator or general strict rules on the language of use in the online setting (Danet & Herring 2007: 19). In some cases, this means choosing the Latin alphabet. Sometimes people prefer to write in the Latin alphabet, even though it is not the official writing system for the language. Some Greek writers see the Latin written Greek as a “challenging, funny, and interesting linguistic innovation” (Tseliga 2007: 134). Writing Greek with the Latin alphabet is more unrestricted since the stress marks and some pedantic grammatical rules are not examined with criticism (Tseliga 2007: 129–130). In addition, using a certain style in online writing also gives the writers the chance to feel that they are part of a group (Silva 2013: 156).

The script can be chosen on the basis of its understandability. The writer can choose a script that is familiar to the recipient so that s/he is able to read it. A situation like this might occur for example when a person wants to chat with a friend who is still in the process of learning the language used in the communication. If the official script is not yet familiar to the recipient, the writing might become readable when it is written in the Latin script. The opposite option is to choose a script that is not familiar to the recipient or is familiar to only a

certain group of recipients. By choosing a certain script the writer can filter the “unwanted” recipients. The Latin script can in some situations be also this unreadable option of a script. This happens in situations where only a certain group of people has learned to write the non-Latin script employing language in the Latin script.

### **2.3.2 Technical reasons**

Writing scripts’ unavailability on keyboards can have some negative effects. The unavailability of the wanted script can result in choosing the Latin script or further, changing the language of communication for example to English instead of one’s mother tongue. In some cases, people might not even try to write their language if the required script is not available (Anderson 2005: 28).

Another technical reason to choose the Latin script is that the official script employs a difficult typing system. This reason was pointed out by the respondents of this research. Arguably most writing systems have been originally designed for writing by hand, not for typing it on electronic devices. This manifests in the form of various problems when it comes to typing some “more exotic” writing systems. Some writing systems have so many graphemes that they cannot fit on a regular keyboard layout so innovative solutions must be made to have all the needed graphemes on the display. In the case of some scripts, the writer needs to remember the visual appearance of the final compound grapheme in order to build it right from pieces of the grapheme. For example, in the case of Cantonese Chinese, the writer needs to know the stroke order of different radicals of the final character in order to write correctly on the Cangjie keyboard (see Figure 1 for the keyboard). This demands skills and time from the writer. If the writing feels too difficult or time consuming, the writer might choose to write in the Latin script.

### **2.3.3 Individual reasons**

In the case of heritage speakers, the use of the Latin script instead of the official non-Latin script indicates that the non-Latin script has not been acquired at all or the scriptal skills have been subject to attrition (Usanova 2019: 117). Therefore, lacking knowledge on writing in the non-Latin alphabet can result to choosing the more familiar Latin script.

Another individual reason for writing in the Latin script would be a personal preference to the script. The preference can be due to several reasons that include life experience, conversations

had, language skills, working environment, personal interests and many other factors. Some of these preferences of the respondents of this research are discussed in more detail in the results (4) and analysing (5) chapters.

#### **2.3.4 Language political reasons**

For writing the non-Latin script languages, the use of the Latin script in instant messaging or online writing is a widespread phenomenon but it is mostly not considered a neutral choice of a script. The Latin alphabet triggers tension concerning local and global identities and the effect of westernisation. On the other hand, the Latin alphabet is favoured in advertisement, restaurant names and in other offline domains. (Spasov 2012: 1488)

Different writing systems can have different political notes to them. For example, according to the Serbian speaking respondents, in Serbia the Cyrillic script might indicate that the writer is politically active or has nationalistic ideas. In Serbian language, the Latin script is more neutral compared to Cyrillic which is mostly used by the Orthodox Church.

If a ruling facet decides to change the writing system for a country or a language, the motives behind the change can be very politically charged. For example, in the case of Croatian, the Latin alphabet was chosen as the official writing system during the disintegration of Yugoslavia, because it emphasised the independence of the language (Spasov 2012: 1488). In the case of Kazakh language, the ongoing script change was triggered by the desire to build a new future and to create distance to the former Russian dominion (Eisenberg 2018). The ideologies, regulations and recommendations created by authorities can affect to the script choices of an individual.

#### **2.4 Societal aspects of script choice**

As mentioned before, this study also takes a look at the script choices that are made by someone else than an individual. Linguistic atmospheres and language policies have an effect on the script choices that are visible in cityscapes, school books and media. In the next subchapters I introduce terms and aspects concerning the script choice that are not controlled by an individual.

### 2.4.1 Digraphia vs. diglossia

Digraphia means that two kinds of scripts that are based on different graphemes are used when writing a language (DeFrancis 1984: 59, Grivelet 2001a: 2). Using two different writing systems for writing different languages by one person is a very common phenomenon, but the true digraphic situation, where speakers communicate in one language but using two different scripts, is less frequent (DeFrancis 1984: 59–60). This kind of true digraphia exist in the case of Serbia (Ivković 2013: 339), Hindi & Urdu (Celine 2017) and Mongolian (Grivelet 2001b). In these languages, digraphia consists of two scripts that are both officially used for writing the languages. Serbian is written in the Latin and Cyrillic scripts. The different usage domains of these scripts in Serbian are discussed in more detail later in this study (Subchapter 5.1.2). Hindi and Urdu are fundamentally the same language, but Hindi is written in the Devanagari script and Urdu in the Nastaliq script. Hindi is mostly used by Indian people and Urdu by Pakistanis and Muslims in India. Mongolian is written in the Cyrillic and Mongolian scripts. The Mongolic script has symbolic and historical value and it is used mostly in religious texts, some newspapers and graffiti written by schoolchildren. The Cyrillic script however is the dominant script and it is used in every domain.

Diglossia is when two co-existing registers of one language are used by the same speakers. The two registers are called “high” and “low” and they indicate to formal and informal usage. These two registers are used in different domains but are still considered as one language. (Coulmas 2003: 229)

Digraphia is the equivalent term of diglossia, the first concerning the written language and the second concerning the spoken language (DeFrancis 1984: 59). They both are phenomena that are about the use of two variants of one language. In the case of digraphia, the terms “high” and “low” registers are not often mentioned when comparing the two different writing systems of a language. An example of this division is in DeFrancis (1984: 64–65) for Mandarin Chinese. DeFrancis argues that the characters are “high” register and pinyin is the “low” register writing system. Otherwise this kind of classification is not a common custom when comparing the scripts in digraphia (Grivelet 2001a: 4). One clear difference between digraphia and diglossia is that in digraphia the choice of script is often done by the writer but in diglossia the speaker rarely can choose whether to use the “high” or “low” variant of speech (Grivelet 2001a: 4–5).

Multiscriptality is a term that is used when two or more scripts are in use by the same language community. Multiscriptality might pose some challenges in language communities but it also gives some freedom to the script users.

In Choksi's research on multiscriptal writing in Santal language in Indian village markets, the multiscriptality affects the social, political and linguistic dynamics. For example, in multiscriptal settings, people do not necessarily interpret the status of different scripts in the same ways as the regimented official language policy would recommend or govern. In addition, at least in the case of Santal in India, the multiscriptal environment and the possibility to use different graphic registers provides a platform to different kinds of expressions. (Choksi 2015: 19–20, 3, 15)

#### **2.4.2 Scripts and language politics**

Writing in the Latin script can be a sociocultural and ideological phenomenon that divides opinions among the speakers and cause language political debates in the media. This has happened in the case of Greeklish: in Greece there are opposing opinions whether Greeklish ruins the historically important Greek script or if Greeklish is an innovative script of the future. (Koutsogiannis 2007: 144, 148–156)

Problems concerning multiscriptality or multilinguality may be very difficult but still possible to solve. Pleasing all the participants in the debates about the choice of the writing script can be challenging. Multiscriptal language communities have some challenges in deciding which of the writing systems is going to be chosen as the official variety. On top of all, the language rights tend to be located low in the scale of priorities of sociocultural rights. This means that the state or the administrative bodies might postpone decisions concerning the writing systems since they are not considered important enough. (Singh 2001: 62, 65, 71)

#### **2.4.3 Linguistic landscape**

Linguistic landscape is a part of multilingualism studies. It investigates the reasons and effects of the multilingual signs by using urban signs and their multilingual written language as material. These signs can be graffiti, billboards, areal maps, emergency guides or any visible text that is on public display. (Backhaus 2006: 1)

For language communities, writing systems provide the means for reflecting the spaces people inhabit, which makes the linguistic landscape studies an important branch of study in sociolinguistics (Choksi 2015: 2). The surrounding signs can indicate the presence of globalisation, migration movements and ethnic identity (Backhaus 2006: 55). In this study, I take a look at the linguistic landscapes briefly from the perspective of the respondents. The surrounding linguistic environment might affect the writing habits of people and inspire them to use the Latin alphabet more if it is also used in the signs.

## **2.5 On the hypotheses**

As it has been revealed in the theoretical background, there are some tendencies in the motives to use the Latin script in languages where it is not the official script. The first hypothesis of this study is that all of the respondents either use the Latin script themselves or that they have at least seen it used by someone else. This hypothesis is motivated by the fact that the Latin script is used globally. Especially the influence of English language and the prevalence of loanwords from English have affected the informal language use in many language communities. The second hypothesis is that the main domain for the Latin script use is online chatting or other Internet mediated communication platforms. This hypothesis stems from the fact that Internet communication is often informal and therefore a very potential domain for using the unofficial writing system. The third hypothesis is that the Latin script use is in some situations the preferred writing system and in some situations the final option. It is presumed that the writers use the Latin script because various reasons.

These hypotheses are supported by the theoretical background of this study and I assume that all three hypotheses are true. In the Chapter 7 of conclusions I return to these research hypotheses and see if the results of this study support the hypotheses as well.

### 3 Methods and Data

In this chapter I introduce the methods that were used for gathering the data. I also present the general information about the data. The data was gathered by the means of the online questionnaire and the interviews. Both of these data gathering methods are explained more in the following chapters alongside with the introduction of data. The inconsistencies and errors that occurred in the methods for collecting data are discussed in the Subchapter 6.1.

#### 3.1 The Online questionnaire

In order to collect data for the study I composed an online questionnaire. It was distributed to my own acquaintances as well as to various Internet platforms. As a theory guideline for composing the questionnaire, I used the material available in KvantiMOTV (Kvantitatiivisten menetelmien tietovaranto, database for quantitative methods).

The questionnaire aimed to be clear and easy to fill, and not too long, since that might make the respondent not want to complete the questionnaire. The respondents and their possible thoughts and answers had to be kept in mind when making the questionnaire. While filling in the questionnaire, the respondent has to feel that the questionnaire is reliable and that her/his answers are not misused. The structure of the questionnaire had to be carefully designed in order to prevent leading questions and therefore distortion in the data. In the beginning of the questionnaire there were general filling instructions along with my personal contact information. The respondents were encouraged to answer without self-censorship by stating that there are no right or wrong answers.

The online questionnaire<sup>2</sup> was composed of multiple-choice questions and several possibilities for the respondent to write more in their own words. I sent the online link of the questionnaire to friends, language and research related Facebook groups and online discussion boards on reddit.com. The online questionnaire was open to fill from 2018-12-19 to 2019-02-03. The platform e-form provided the means to convert the gathered data to a spreadsheet file. From this file, I could see the quantitative data that was gathered with the multiple-choice

---

<sup>2</sup> The questionnaire is available for a closer look in Attachment 1. The questionnaire platform “e-lomake” was provided by the University of Helsinki.

questions as well as the qualitative data of open replies. With spreadsheet functions I got quantitative data that was simple enough to process and analyse.<sup>3</sup>

The first nine questions in the questionnaire were about the background and the language of the respondent. The respondents told about their age, gender, occupation, nationality and country of residence. These questions were obligatory, and they provided me with metadata. Next in the questionnaire it was asked whether the respondent uses the Latin script when writing her or his language. If the respondent answered “yes” s/he was asked to answer to the following 10 questions about how often the respondent writes in the Latin script, what kind of style is used, what are the reasons for the Latin script use, what are the domains of use and does the respondent like to write in the Latin script. If the respondent answered that s/he does not use the Latin script when writing her/his language, s/he could proceed to a question that asked whether the respondent likes to write her/his language in the official script.

The next question was about whether the respondent has seen her/his language written in the Latin script by others. For those who answered “yes” there were two additional questions about the domains where s/he had seen it being used and the things that s/he had seen written in the Latin script. If the respondent answered “no” to the question about the use of the Latin script by others, s/he could proceed straight to the five last questions. They concerned the use of different scripts on different electronic devices. The respondents were also asked about their opinions on the functionality of these devices concerning the writing on different scripts.

Because the number of respondents was not set beforehand, it was important to collect as much data as possible from the respondents that did the questionnaire. Therefore, the online questionnaire contained abundance of questions about all kinds of factors that might contribute to the results of the data. Mainly I wanted to find out if the respondents use the Latin script when writing their language or if they have seen it used somewhere by others. Another thing that I was interested in was the functionality of different electronic devices when dealing with the Latin and non-Latin keyboards.

Since the aim was to get as many replies to the questionnaire as possible, many of the questions were in a multiple-choice format which enabled that the data would be as easy and quick to analyse as possible. Additionally, in the end of every question there was a chance for the respondents to write more about their thoughts in their own words. There was either an

---

<sup>3</sup> See Attachment 2 for examples of filled forms.



option “other” and a chance to specify this option shortly, or a separate question in the style of “here you can write more about your answer in 17a”. These kinds of open questions had a larger area for typing the answer so that the respondent would feel that s/he can write as much as s/he wants. With the open answer questions the respondents could freely express their thoughts about the given questions. In the end of the questionnaire, there was one more chance to write about any thoughts that the respondent might have gotten while doing the questionnaire. The open phrases written by the respondents are used in this research as qualitative data, which is analysed alongside with the questionnaire’s quantitative data and the data gotten from the interviews.

I formed the list of multiple choices myself and the possible answers were inspired by other studies about the use of the Latin script (for example Hämeen-Anttila 2016, Grivelet 2001a). I also got some tips and ideas for possible questions from my pilot study where eight people filled in the online questionnaire and gave their comments about the process. The questions in the final questionnaire dealt with personal Latin script use of the respondent and the Latin script use by others, as perceived from the point of view of the respondents. By others I mean either individual people that the respondent knows (friends, family members, colleagues etc.) or some organisation or authority (media, press, advertisement, road signs etc.).

### **3.2 The Interviews**

Another method that I used for collecting data was to have interviews with four respondents. For the theoretical and methodological background, I used the Hirsijärvi & Hurme (2000) guide for scientific interviews. As the topic of the interviews was very personal, certain ethical manners had to be kept in mind throughout the whole process (more on the research ethics in Subchapter 3.4). The type of interview that was used in this study is a half-structured interview (Hirsijärvi & Hurme 2000: 37). The interview was partly planned since I had thought of some questions beforehand but apart from those questions, the interview was carried out spontaneously in a conversational way. With the interviews I got more answers to questions that have a personal tone to them. The interviews gave me a chance to discuss the subject from different perspectives and have an interactive discussion with the respondents. I got information for example on what the respondents feels about using the Latin script and how exactly do they use the keyboard on their phone between two or more writing systems.

The interviewees did the online questionnaire beforehand so that I had some data to work on before the interview. This way their answers are also a part of the quantitative online questionnaire database.

I interviewed four people with different linguistic, social and geographical backgrounds. Therefore these four interviewees form a very heterogeneous group and they do not represent all the other respondents who were not interviewed. The interviewees are my own acquaintances that I have gotten to know over the years. I did not follow all of the guidelines in Hirsijärvi & Hurme (2000: 72) where using own acquaintances as respondents is not recommended. The already existing relations with the interviewees might compromise the objectivity of the research and make the interviewees consciously or unconsciously modify their answers. Nonetheless, I asked these acquaintances of mine to be interviewees because they use very different writing systems when writing their language. The respondents speak Russian, Farsi, Cantonese Chinese and Thai as their mother tongue. In addition, I knew that they were easy to reach and willing to participate in the research, which made the interviews eloquent and interesting for both the interviewee and the interviewer.

Two of the interviews were held in person in a public cafeteria and with the other two we had the interview via Whatsapp call. I recorded all of the interviews with a recorder and then transcribed the interviews to written form. The interviews took 30–50min. In the interviews we went through some of the answers that the respondent had given in the online questionnaire since one of the goals of the interviews was to fill in the gaps that were left behind from the questionnaire. I then asked some additional questions to get some insight on the sociolinguistic atmospheres in these four language communities. The additional questions that I had planned beforehand to ask the interviewees were for example about the language political situation in their childhood country of residence. I also asked whether older people in their language community would understand their language if it was written in the Latin script and if abbreviations such as NASA, AIDS or USA are written in the Latin script or in the official script of the language in question. Additionally, I wanted to know more about the general phone keyboard usage of the respondents, so I asked them questions about the ways they type on keyboards on electronic devices and what kind of possible problems are related to the typing. The interviews turned out to be so loosely structured and spontaneous that I cannot provide a list of all the questions that were asked during the interviews.

### 3.3 The Data

In this chapter I introduce the data that I gathered for this study. First, I introduce the quantitative data that was gathered with the online questionnaire. Then I shortly present the qualitative data which consists of the open question answers of the questionnaire and the four interviews. The whole data was not possible to attach to the end of this paper due to the limited space of text format and the large amount of data. However, four randomly selected online questionnaire datasets are in the attachments to give an idea of what kind of data was processed for this study.

Overall 150 datasets were submitted by the means of the online questionnaire. Out of these, eight datasets had to be taken out because they were not suitable for the research. One of these eight datasets was done by an underage person and one dataset was about the German language which is written in the Latin alphabet and therefore the dataset was not suitable for this study. The rest six datasets were deleted because they were duplicates probably due to technical issues occurring during the submission of the form.

All together 142 datasets could be taken as data for the research. The answers of the respondents represent 32 languages and language variants that utilise 22 different writing systems. They live in 35 countries and represent 37 nationalities or ethnicities which include double nationalities and undefined identifications of respondents. 94 respondents are women and 48 are men. The age range of the respondents is 18–70. Majority of them (111) are in their 20's or 30's with four respondents under 20 years old and 27 respondents over 40 years old. Education level of the respondents range from high school and vocational school students to university students and doctors of Philosophy. Most of the respondents are university graduates (97) or university students (27). The respondents work in 24 different fields of profession. Since the online questionnaire was in English and the Internet accessibility was required to fill it in, the respondents quite likely were well educated and lived in a place with high infrastructure.

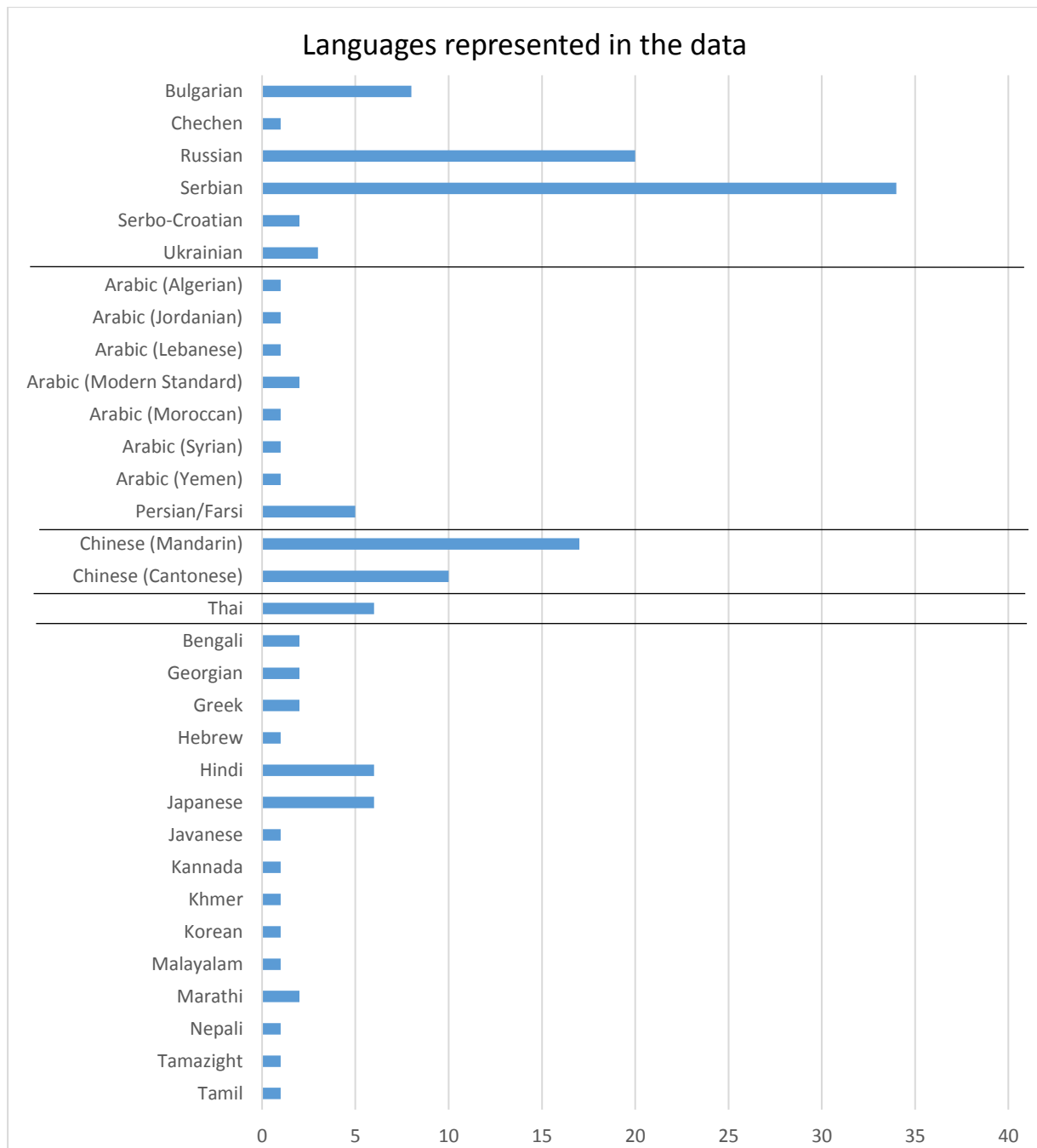
As mentioned in the introduction chapter, I have divided the 142 respondents into five groups according to the official script of their language. The five groups of scripts are the following: “Cyrillic”, “Arabic”, “Chinese” (including both traditional and simplified characters), “Thai” and “Others”. The first four writing systems are those that the interviewees use when writing their language. In the group “Others” are languages like Japanese, Greek, Georgian, Javanese,

Hindi and Marathi. Although Japanese uses the Chinese characters in its writing system, in this study, Japanese is treated separately from the Chinese languages. This is because there is a geographical and linguistic separation between Chinese and Japanese languages.

Additionally, Japanese utilises a three-script-system of Hiragana, Katakana and Kanji (adopted Chinese characters) whereas Chinese uses only the Chinese characters.

As is visible in Figure 2, most of the languages are represented by only 1–6 people. Serbian, Russian and Mandarin Chinese have the most representatives. Since the data is divided into groups, the less represented languages are treated as a part of a bigger group.

The “Cyrillic” group consists of Russian (20), Bulgarian (8), Ukrainian (3), Chechen (1), Serbian (33) and Serbo-Croatian (2) speakers. Later in the text, I discuss the data of the Serbo-Croatian speakers alongside with the data of the Serbian speakers. According to the data, the two Serbo-Croatian speakers use both the Latin and Cyrillic scripts actively, which indicates that the Serbo-Croatian respondents are influenced by the digraphic situation in Serbia. Therefore, for the purposes of this study, discussing these two language variants together is reasonable. In the “Arabic” group, the following languages and language variants are included: Farsi/Persian (5), Yemeni Arabic (1), Syrian Arabic (1), Moroccan Arabic (1), Modern standard Arabic (2), Lebanese Arabic (1), Jordanian Arabic (1) and Algerian Arabic (1). The “Chinese” group consists of two languages: Mandarin Chinese (17) and Cantonese Chinese (10). Two of the Mandarin Chinese speakers are bilingual. In the analysing Chapter 5 this group is further divided into Taiwan Mandarin, mainland China Mandarin and Hong Kong Cantonese according to the different writing conventions of the characters. The “Thai” group includes only the Thai language speakers (6). The last group “Others” is a mixture of the rest of the languages that are not included in the four previous groups. In the group, Japanese is represented by six respondents. In the same group, there are six Hindi speakers, out of which one is bilingual. The other languages in this group are Tamil (1), Tamazight (1), Nepali (1), Marathi (2), Malayalam (1), Korean (1), Khmer (1), Kannada (1), Javanese (1), Hebrew (1), Greek (2), Georgian (2) and Bengali (2).



*Figure 2. The different languages and language variants that are spoken by the respondents. The black horizontal lines indicate the five separate groups.*

Especially the “Thai” group is small in quantity of respondents and consists of the same number of speakers as the speakers of Hindi and Japanese in the “Others” group. Therefore, any following generalisations regarding the “Thai” group are as valid as the generalisations made about the languages that are represented by the same number of respondents. The group division is made and contained throughout this research strictly because of the scripts that the interviewees use for writing their languages. Therefore, even though Thai language has a

relatively small number of representatives, it is treated as its own group since one of the interviewees is a representative of Thai writers.

In the results Chapter 4, I look into the results of the data and introduce it in numbers, tables and figures. The data is introduced in its entirety, but the numbers are presented also according to the five groups of writing systems.

In addition to the quantitative data, the online questionnaire provided good amount of freely composed sentences from the respondents. All together the data contains 481 written answers of different lengths. Some of them are written out of lack of the preferred option in the multiple-choice questions. The rest of the open answers are general comments on the topic or specifications to the multiple-choice question answers. I introduce some of the results of these open question answers in Chapter 4 alongside with the quantitative data of the online questionnaire. In the analysis Chapter 5, I quote some of the open question answers, as they are very interesting and provide information straight from the respondent. They are presented in their original forms to prevent possible inaccurate corrections and in order to preserve the personal voice of the respondent.

Four interviews consist of roughly over two hours of semi-structured discussion between me and the interviewee. I introduce the results of the interviews separately in subchapter 4.5 in a narrative text form. Further in the text in the analysing Chapter 5, I compare the interviews, the quantitative data and answers to open questions. In order to compare, I searched for similarities and tendencies as well as any other explicit differences between the datasets.

### **3.4 Research ethics**

There are some ethical issues that need to be considered in studies where people share their personal information. Some of the questions asked in the questionnaire or during the interview might have felt intruding or difficult to answer. Sharing one's personal opinion on politically and nationally related language questions can be a delicate issue that might cause the respondent some discomfort. However, in this study I followed the informed consent principles. This means that all the respondents were asked for their consent to using the information they gave as data for research purposes. The same consent was implemented for the four interviews, since the interviewees had filled out the questionnaire before the interviews. Additionally, in the questionnaire, the respondents were provided the possibility to

comment freely both their own answers and the research in general. The respondents of the questionnaire were all informed that their information is collected in anonymous manner since no names, phone numbers, addresses or such were asked in the questionnaire. The interviewees are also referred to without names or other information that might compromise the anonymity. Before giving their answers, all the respondents were informed about the general content of the questionnaire or the interview so they had an understanding of what kind of questions they will be asked.

During the process of this study I have kept in mind that languages are spoken in different political situations and that nationality has a big role in the language identity formation. However, in this study the focus of discussion is in the individual and technological side of things. The political and national regimes might have an impact on the results of this study, but the possible effects are not discussed in detail.

## 4 Results

In this chapter I introduce the most relevant results of the data. First, I go through the online questionnaire datasets. I look into the quantitative data as well as some of the most relevant open question answers. After these I present the results of the interviews.

### 4.1 Online questionnaire results

Out of the 142 respondents 91 say that they use the Latin alphabet when writing their language. That means that less than half of the respondents (51) do not use the Latin alphabet when writing their language. The 91 people who informed that they use the Latin script, answered to some additional questions about their use of the Latin script. The results to the additional questions are discussed in the following subchapters. While filling in the questionnaire, some respondents answered questions that they were requested to skip. This means that the data has some inconsistencies. However, all available data is considered here and in the whole study, regardless the possible contradictions in the answers to separate questions.

#### 4.1.1 Use of the Latin script by the respondent

In the “Cyrillic” group, the majority say that they use the Latin script but in the other groups the answers to this are close to half-and-half, the “yes” answers still being more frequent (see Table 4 for more details). Of the all the respondents who use the Latin script, 13 use it very seldom, 14 use it at least once a month, 25 use it at least once a week and 39 use it every day. The replies concerning different usage frequencies are distributed quite evenly between groups. The “Cyrillic” group was a notable exception to this, as most of the Serbian speakers use the Latin script every day.

Majority of the ones who use the Latin alphabet had started using it during primary school years. 53 people had started at the age of 4–7 or 8–11. The second biggest group with 28 people is the ones who started using the Latin alphabet when they were teenagers, 12–15 or 16–19 years old. Seven people started when they were over 20, one started at over 30 and one at over 40. One person did not know when he started using the Latin script. These replies are also evenly distributed between the different groups of scripts. Only the speakers of Farsi have all started to use the Latin script when they were 12–15 years old. All the speakers of other languages have chosen various answers.



In the questionnaire there was not a question on which year the respondent has started to use the Latin script. I calculated the year by comparing the current ages of the respondents with the ages when they began to use the Latin script. With this I wanted to see if the use of Latin script has become more common among a language community at some specific point in time. If many respondents using the same script or speaking the same language would have started to use the Latin script around the same time, some assumptions on the situation in sociolinguistics or technological development could be made.

The outcome of this experiment was not very fruitful since no clear connection between the non-Latin scripts and the start of the Latin script use was found. The most potential finding is the relatively narrow timespan of Latin script use starting point by the speakers of the languages that are indigenous in India and Bangladesh. The six respondents who use the Latin script when writing these languages, have started using it in average between 2005 and 2015. One Kannada speaker, in addition to these six, has started to use the Latin script in 1995 but this early usage might be explained by the respondent's interest in languages since he is a doctor in linguistics. In other language and script groups the respondents had started to use the Latin script during even a longer period of time than ten years. Therefore, even though ten years is a wide timespan, it is still the narrowest timespan that was found in the data.

If these results were hypothetically relevant, an assumption could be made that online and instant messaging in the Latin script gained popularity in India and Bangladesh between 2005 and 2015. If some language group or script group had started to use the Latin script at the same time and a long time ago, a hypothetical assumption could have been made that the script for this/these languages was not yet available on keyboards or included in the Unicode.

The respondents replied to a question about the length of text that they write in the Latin alphabet. Out of 91 respondents, three indicate in their answers that they write just some letters within a word in the Latin alphabet, whereas 19 say that they write some words inside a sentence in the Latin alphabet. Six respondents write complete sentences and eight write several sentences within a text in the Latin script. 24 of the respondents say that the length of the text depends on what s/he is about to write and where. Over a third of the respondents (32) write the whole text in the Latin script. Out of these, 21 respondents are Serbian or Serbo-Croatian speakers, who are influenced by the digraphic environment. According to three Serbian respondents in the open answers, in Serbian language one should not mix the Cyrillic

and Latin scripts. Therefore in Serbian, it is possible to write anything in the Latin or the Cyrillic script as long as one of them is chosen as the only script used within the text.

91 respondents answered to a question whether they use the Latin alphabet in handwriting or not. Out of these, 48 people do write the Latin alphabet by hand and 43 people do not. The result of this question is therefore almost equally divided for both options. In the “Cyrillic” group, both replies have almost the same amount of answers, but Russian and Bulgarian speakers have the majority of “no” while Ukrainian, Serbian, Serbo-Croatian and Chechen have the majority of “yes”. In the “Chinese” and “Thai” groups, a clear majority of respondents write the Latin script by hand. In the “Arabic” group only one writes the Latin script by hand in his language. In the group of “Others” the majority of respondents do not use the Latin script when writing by hand. In this group, the ones who have replied “yes” (6) are representatives of Japanese and Hindi. In the open answers respondents specified that they write the Latin script by hand if they need to write some quick notes, shopping lists or addresses to international packages. Some mentioned that they need to write in the Latin script by hand out of professional requirements.

92 respondents answered to a question about whether the respondent uses some official spelling rules when writing their language in the Latin script (for example standardised romanisation rules) and the result is that 56 people do use spelling rules, 22 do not and 14 do not know. A bit over a half of the respondents therefore use some spelling rules. The number of respondents who did not know how to answer to this question was quite high. This can also imply that the question was unclear. 91 respondents answered to a question on the style of language they use when they write their language with the Latin alphabet. “Formal standard language” has 14 replies, “vernacular/spoken language” has 23 replies and “both” has over a half of the replies with 54 answers.

In the “Cyrillic” group, seven respondents do not know if they use the official spelling rules, but the majority of the rest do use them. Most of the Serbian and Serbo-Croatian speakers use both standard and spoken language when writing the Latin script. The Russian speakers use either both or just standard language and the Bulgarian speakers either both or spoken language. The answers of the Arabic script users are diverse when it comes to the following of spelling rules but with only one exception all the Arabic script users write the Latin script in the spoken language. In the open question answers one Farsi speaking respondent mentions an interesting detail: if a text is written in the Latin script in Farsi, the respondent is able to

determine from what region in Iran the writer comes from. This is because the spoken language reveals the regional dialect of the writer. It seems that the Latin script is more suitable for the unofficial messages than the Arabic script. The respondents of the “Chinese” group mostly use some spelling rules. In addition, they mostly use both formal and vernacular language or just formal language when writing in the Latin script. In the open question answers two of the Mandarin Chinese speakers have specified that it sometimes depends on the function of the text, whether the spelling rules should be used or not. All the three Thai respondents who use the Latin script have different answers “yes”, “no” and “I don’t know” to the question about following spelling rules. Two of them use both formal and vernacular language and one uses only vernacular language. In the group “Others” all the Japanese respondents follow official spelling rules and write the Latin script in both formal and vernacular language. On the contrary the speakers of languages spoken in India and Bangladesh do not follow official spelling rules and they use only spoken language when writing in the Latin script. In the open question answers two Hindi speaking respondents specify that even though there are no official spelling rules, everybody uses an informal standard for writing Hindi in the Latin script. One Hindi speaker does not know if she follows the spelling rules or not.

The questionnaire collected 91 answers to a question whether the respondents write their language in the Latin alphabet when writing official papers. Almost two-thirds (61) answered “no” to this question and the remaining 30 respondents answered “yes”. The “yes” replies are distributed between several languages, 15 of them being Serbian and Serbo-Croatian speakers.

In the questionnaire, there was a question if the respondent likes to write her/his language in the Latin script and another question asking if s/he likes to write her/his language in the official script. Only the respondents who use the Latin script when writing their language replied to the former question. Out of 94 people, 42 say that they like writing in the Latin alphabet and 45 say that they do not like writing with it. Therefore “yes” and “no” answers are distributed quite evenly. However, observing per group, the distribution of replies is not so even. In the “Cyrillic” and “Chinese” groups the majority replied “yes” while in the “Arabic” and “Others” group the majority replied “no”. The replies in the “Thai” group are evenly distributed.

The question if the respondent likes to write her/his language in the official script got 141 replies, since all the respondents were allowed to answer this question. One person skipped this question. Out of all the respondents, 116 like to write their language in the official writing system and only 20 people do not like writing their language with it. The greatest contrast between “yes” and “no” answers are in the “Cyrillic” and “Others” groups where the respondents are not unanimous. The “Chinese” group also has a notably large difference in the distribution of replies. The majority of each group chose “yes”.

Seven respondents say that they are indifferent about writing their language in the Latin script and six say that they are indifferent about writing in the non-Latin script.

<b>Do you like writing in the Latin script?</b>	<b>All</b>	<b>Cyrillic</b>	<b>Arabic</b>	<b>Chinese</b>	<b>Thai</b>	<b>Others</b>
yes	42	25	2	9	1	5
no	44	22	6	4	1	11
indifferent	7	3	0	2	1	1
<b>all</b>	<b>93</b>	<b>50</b>	<b>8</b>	<b>15</b>	<b>3</b>	<b>17</b>

<b>Do you like writing in the official non-Latin script?</b>	<b>All</b>	<b>Cyrillic</b>	<b>Arabic</b>	<b>Chinese</b>	<b>Thai</b>	<b>Others</b>
yes	115	60	9	19	3	24
no	20	4	4	6	2	4
indifferent	6	2	0	2	1	1
<b>all</b>	<b>141</b>	<b>66</b>	<b>13</b>	<b>27</b>	<b>6</b>	<b>29</b>

*Table 1. The distribution of replies in the different groups concerning the opinions on writing with different writing systems.*

The contrast between these opinions is distinct. The replies to the question about writing in the Latin script are almost equally divided into two but over three-fourths like writing their language using the official writing system. In the open question answers the respondents have mentioned many pros and cons concerning the Latin script and the official script. Some respondents admit that the Latin script is sometimes the most convenient script to use but some say that the Latin script looks boring or that it does not suit the language. Some respondents say that the official script is difficult to type but on the other hand the official script is mentioned to have cultural or historical importance.

#### 4.1.1.1 The domains of use

In the question about the domains of the Latin script use, one person skipped the multiple-choice options. In addition, three people who had answered “no” to using the Latin alphabet in their language, answered erroneously to the question about the domains of use.

<b>Latin script usage domains</b>	<b>All (95)</b>	<b>Cyrillic (52)</b>	<b>Arabic (8)</b>	<b>Chinese (15)</b>	<b>Thai (3)</b>	<b>Other (17)</b>
chatting with friends	70	38	8	9	3	12
chatting with family members	50	29	5	5	2	9
chatting with strangers on the internet	41	27	2	5	2	5
posting on social media	50	32	2	6	3	7
at work assignments	35	24	0	7	3	1
at school assignments	21	10	0	5	3	3
writing to foreign friends	56	37	3	8	2	5
when translating my language to some other language	51	30	4	9	2	6
writing info about myself when visiting abroad (for example name and address at hotels)	57	39	2	6	3	7
I don't know/remember	0	0	0	0	0	0
other	16	12	0	1	0	3

Table 2. The distribution of replies concerning the Latin script usage domains.

16 people use the Latin script in contexts that were not listed in the questionnaire. The respondents have specified in the open question answers that the Latin script is used in general when writing on a computer or when searching something on the Internet. The respondents use it when doing online shopping, writing an SMS or doing international business. The Latin script is also used when teaching a language and when translating. Some people use Latin alphabet in exams.

Using the Latin script when chatting with friends is one of the most popular domains by all of the five groups. The Latin script is used the least at school and in work assignments. The “Cyrillic” group gave a lot more answers than the other groups. This is because the Serbian and Serbo-Croatian speaking respondents actively use the Latin script in many domains due to the digraphic environment in Serbia. For example, out of the 57 answers to “writing info about myself when visiting abroad”, 27 answers are given by Serbian and Serbo-Croatian speakers, which is nearly half of all the answers.

#### 4.1.1.2 The reasons of use

There was some incoherence in the results concerning the question about the reasons to use the Latin script. One person who said that she uses the Latin script, did not answer to this question and two people who answered that they do not use the Latin alphabet when writing their language, gave their opinions to this question.

Reasons such as “shorter binary code”, “just for fun”, “others do it too” and “using western alphabet is cool” are the least popular reasons. From these results it could be said that the direct influence of other people is not very strong when it comes to choosing a script. The rest of the reasons got relatively equal amount of replies, including the option “other”.

Reasons for using the Latin script	All (92)	Cyrillic (49)	Arabic (8)	Chinese (15)	Thai (3)	Others (17)
it feels easier than the official writing	32	14	3	5	3	7
it is faster than the official writing	33	15	4	7	2	5
the binary code of Latin letters is shorter	12	4	1	5	2	0
other keyboards were not available	33	15	6	4	0	8
just for fun	17	9	3	1	1	3
others do it too	9	2	3	0	1	3
using western alphabet is cool	1	0	0	0	1	0
some things are not translated to my language	30	11	5	6	3	5
other	31	21	0	4	0	6

Table 3. The distribution of replies concerning the reasons for using the Latin alphabet.

Additionally, in the open question answers 31 respondents thought of other reasons to use the Latin script. In the case of the Mandarin Chinese speaking respondents, they use the Latin script because they use the Pinyin input system to write the Chinese characters. Another reason mentioned is that if they need to write something online that is blocked by the Chinese government, the Latin alphabet can be used in order to avoid censorship. The Cantonese speakers use the Latin alphabet if they forget how to write the character correctly. According to one Japanese speaker the Latin script is also used when some respondents want to write “unofficial” loanwords that do not yet have a writing convention in the Japanese writing systems. One Hindi speaking respondent tells that the Latin script used to be the only script available for different electronic devices, so she had to use the Latin script in order to communicate. When the Devanagari keyboards became available, the respondent did not start learning the new way to write so she continued with the Latin writing system.

Some Serbian and Serbo-Croatian speakers use the Latin script out of habit. Another reason for some respondents to use the Latin script is to get more numerous search results on the Internet. Also, tagging friends or locations on social media, might only work with the Latin alphabet. For some, laziness to switch the keyboard is the reason to use the Latin script. Some respondents mention that they use the Latin script when chatting because it is more suitable for writing the spoken language.

#### 4.1.2 Use of the Latin script by others

Out of the 142 respondents, 131 have seen their language written somewhere in the Latin script. Two of these respondents did not give answers to where they had seen it used and four of these did not say what they had seen written in the Latin alphabet in their language. One person says that she has not seen her language written in the Latin alphabet, but she gave her answers to what is written in the Latin script anyway.

Uses the Latin script	Has seen the Latin script used by others	All (142)	Cyrillic (67)	Arabic (13)	Chinese (27)	Thai (6)	Others (29)
Yes	Yes	89	49	8	13	3	16
No	Yes	42	18	5	5	3	11
Yes	No	2	0	0	2	0	0
No	No	9	0	0	7	0	2

*Table 4. The distribution of replies in the questions concerning the respondent's individual use of the Latin script and the use of the Latin script by others.*

There is a greater number of respondent who have seen their language written in the Latin alphabet than those who have used the Latin alphabet themselves. In every group the “yes” responses are more frequent to the question about seeing others using the Latin script than to the question about the respondents using the Latin script themselves. As mentioned in the beginning of Subchapter 4.1.1, a clear majority in the “Cyrillic” group uses the Latin script whereas in the other groups the replies are distributed quite evenly between the “yes” and “no” options. When asked whether the respondent has seen others use the Latin script, the replies are more distinctly towards the “yes” option. All of the respondents from the “Cyrillic”, “Arabic” and “Thai” group say that they have seen it used by others. In the “Chinese” and “Others” groups a clear majority has seen the Latin script used by others but not all of the respondents in these groups use the Latin script themselves.

Most of the respondents use the Latin script themselves and have seen it used by others too. Less than a half of all the respondents say that they do not use the Latin script themselves but have seen it used by others. The replies for these two options are quite evenly distributed between the groups with respect to the amounts of respondents in the groups. Two respondents from the “Chinese” group say that they use the Latin script themselves but have not seen it used by others. Nine respondents, two from the “Others” and seven from the “Chinese” groups say that they do not use the Latin script themselves, nor have they seen it used by others when writing their languages. These nine respondents represent three languages: Mandarin Chinese, Cantonese Chinese and Marathi.

#### 4.1.2.1 The domains of use by others

Domains of use	All (129)	Cyrillic (65)	Arabic (13)	Chinese (18)	Thai (6)	Others (27)
friends chatting with electronic devices	101	50	13	11	6	21
family members chatting with electronic devices	72	38	9	7	2	16
strangers on the internet chatting	92	53	11	7	3	18
social media posts	92	47	9	10	6	20
Latin alphabet is used at work	38	27	0	5	2	4
on street signs	66	38	4	11	4	9
in TV advertisement	62	37	2	10	3	10
in newspaper advertisement	54	36	2	7	2	7
in shop products	65	40	4	10	2	9
poetry	38	31	0	4	0	3
novels	39	31	0	6	0	2
religious texts	11	8	0	2	0	1
children's books	31	21	1	6	1	2
language learning materials	65	38	3	14	4	6
hand written texts	63	36	4	10	3	10
other	15	8	2	2	0	3

Table 5. The distribution of replies concerning the domains of use by others from the perception of the respondent.

129 gave their replies to the multiple-choice question about the domains where they have seen the Latin script used by others. The most popular domains are those that are related to Internet and electronic devices. The least popular domains are religious texts and children's books. In the open question answers 16 respondents inform that they have seen the Latin alphabet being used in some other places that were not listed in the questionnaire. One of them is the words that have been adapted to English, like *matreshka* from Russian and *anime*, *sushi* and *tsunami* from Japanese. Other places where the Latin alphabet is used is the business names, linguistic



papers, addresses and identification documents. Also, T-shirts, mugs, comics, animations, postcards, posters, subtitles and song lyrics have sometimes text written in the Latin alphabet in the language of the respondent.

Between the groups, the respondents of “Cyrillic” represent a clear majority in many of the options. The majority of replies are from the “Cyrillic” group, especially concerning the options “using the Latin alphabet at work”, “in newspaper advertisement”, “poetry”, “novels” and “children’s books”. The other replies are distributed quite evenly between groups concerning the amounts of respondents in each group.

#### 4.1.3 What is written in the Latin script?

In this chapter I look into the results of the question concerning the nature of words that are written in the Latin script. Out of the multiple-choices the respondents chose what they themselves write as well as what they have seen written by others. To this question, 128 respondents gave their answers, which means that four respondent skipped this part in the questionnaire although they had seen the Latin alphabet being used somewhere in their language. One person answered to this question even though he had answered earlier that he has not seen the Latin script used by others.

<b>What is written in the Latin script?</b>	<b>All (128)</b>	<b>Cyrillic (65)</b>	<b>Arabic (13)</b>	<b>Chinese (18)</b>	<b>Thai (6)</b>	<b>Other (28)</b>
words from English	61	22	6	13	6	14
words from other foreign languages	39	14	5	6	5	9
foreign names of people	43	12	3	11	6	11
foreign names of places/buildings/other geographical names	39	10	4	8	5	12
pronunciation rules for your language	25	8	4	8	1	4
abbreviations	32	11	6	8	2	5
anything is possible to write in the Latin alphabet in your language	67	39	9	4	3	12
other	10	5	0	0	0	5

*Table 6. The distribution of replies on the question what they write or what they have seen written in the Latin script.*

“Words from English” and “anything is possible to write in the Latin script” are the most popular choices. In choosing the latter option, the “Cyrillic” group is a clear majority. The replies to the latter option give some interesting insight about the possible suitability of the Latin script for writing the languages presented in the data. Some of the respondents who

have chosen this option are not the same respondents who use the Latin script themselves. A bit over a half of all the respondents say that anything is possible to write in the Latin script in their language. Most of the replies are naturally from the “Cyrillic” group and more specifically from the Serbian and Serbo-Croatian speakers. The fewest replies for this option are from the “Thai” and “Chinese” groups. Within the “Chinese” group, four of the mainland China Mandarin speakers and only one Cantonese Chinese speaker has chosen this, resulting as zero replies from the Taiwan Mandarin Chinese speakers. Around half of the respondents in the “Arabic” and “Others” groups say that the Latin script can be used to write anything. In the “Others” group many of the replies are from speakers of languages spoken in India and Bangladesh. On the contrary, only one Japanese says that the Latin alphabet can be used to write anything in Japanese.

In the open question answers 10 of all the respondents say that there are other things that can be written in the Latin script that were not introduced in the questionnaire. These things are brand names, names of media resources, e-mail addresses and websites. People also use the Latin script to write science terminology, names of movies and songs and to transcribe words and names from the “original” language.

#### **4.1.4 Writing on electronic devices**

When it is required or preferred to write with more than one writing systems, mobile phones need to have a function to switch between different keyboards. The respondents answered to a question about the functionality of switching between different keyboards on their phone.

The keyboards are discussed here by their operating systems which often dictate the default keyboard layout and keyboard functions on a mobile phone. Some of the respondents might have downloaded other keyboard applications onto their phones according to their own preferences. In addition, operating system versions differ from phone to phone which might result to some differences in the keyboards. Despite all of these possible variables, in this study I consider the operating systems as if all the phones with the same operating system employ the same kind of keyboard.

Because of methodological reasons, I asked the respondents to indicate their phone brand instead of the operating system. Some respondents might have not known the operating system of their phone which might have resulted as false data. The respondents use 22 different phone brands out of which Apple, Huawei, Samsung and Xiaomi are the most

represented. Out of the 141 respondents, who have a phone, 84 use an Android operating system and 55 uses an iOS. Two respondents have phones that have neither of these operating systems. These two phones are not included in Figure 4.

Out of the 141 respondents 64 say that switching between keyboards works very well and 44 say that it works pretty well. Therefore, over three fourths are generally satisfied with this function. According to 12 respondents switching works quite poorly and two people thinks it works really badly. These two speakers, a Serbian speaker and a Tamazight speaker, both have an Android phone. For 11 people the functionality varies depending on the situation and eight people do not know what to answer to this question.

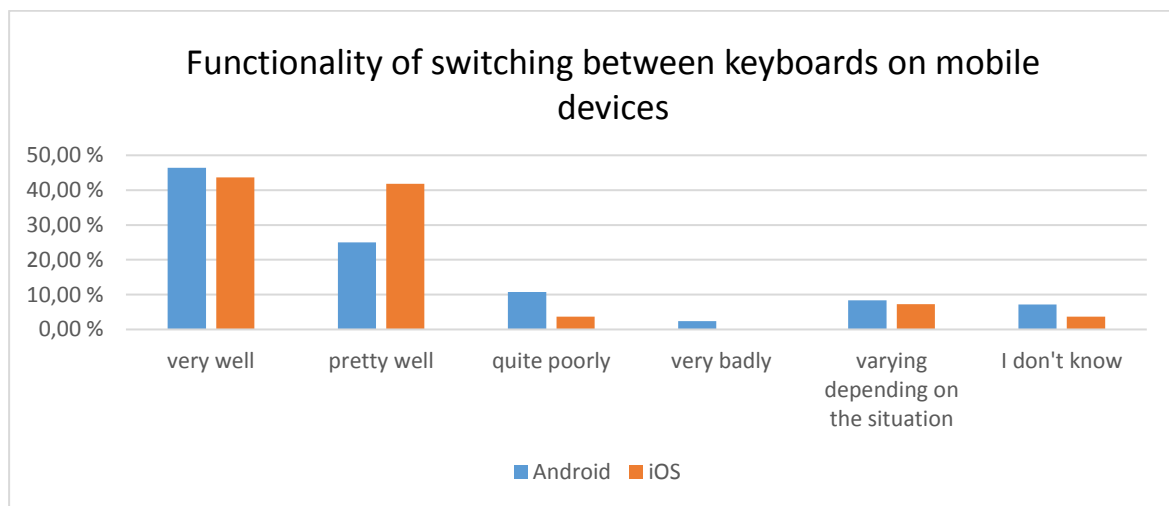


Figure 4. The distribution of keyboard functionality on the two phone operating systems that are most used by the respondents.

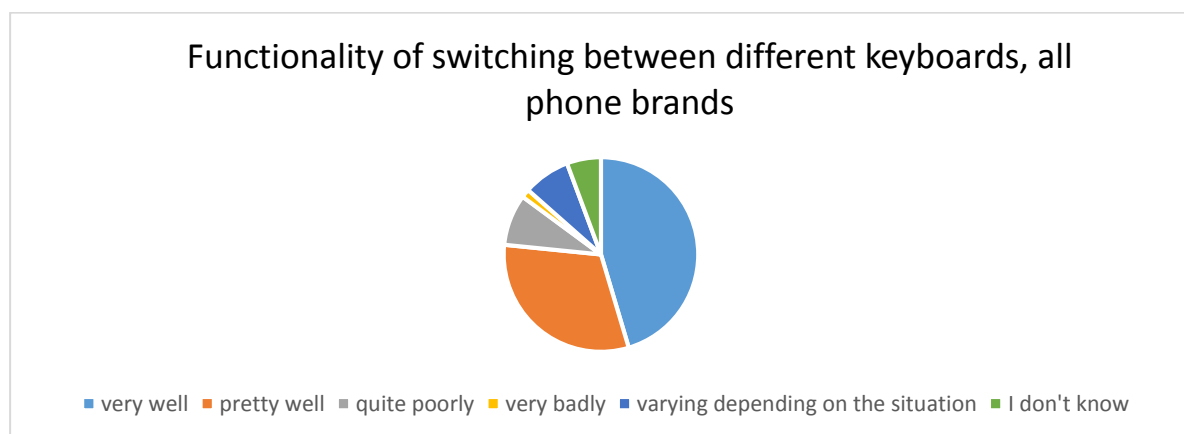
All the five groups have users of both Android and iOS phones. In the “Cyrillic” group four respondents think switching between keyboards works quite poorly and three of these use an Android phone. The other replies in this group are distributed unpredictably between the operating systems. Three out of all the 13 “Arabic” group respondents have chosen “quite poorly” as their answer for this question which is 23% of all the replies in the group.

However, these replies represent both Android and iOS phones so no generalisations can be made. In the “Chinese” group only one Cantonese speaker says switching between keyboards works quite poorly. The other respondents of the group are satisfied with the functionality on their phones. In the “Thai” group all of the respondents are satisfied as well. There is variation in the answers of the “Others” group because of the nature of the group. All of the different

options for usability have gotten replies from this group but over a half think their phones work either very well or pretty well when switching between keyboards.

In the “Others” group, one person answered “I don’t know” to both questions concerning the usability of electronic devices since he does not write his language (Nepali) on his phone. In the open question answers he specified that he is not even sure if the Devanagari system for his phone was available. Additionally, one Javanese speaker and one Tamazight speaker have answered that a keyboard that would function well is difficult to find. One Marathi speaker says that the Devanagari script is difficult to type and that writing in English is much easier. Some people have made their own keyboard layouts, or they have had to install new keyboards since a suitable keyboard has not been available on their phones automatically.

In the questionnaire there was also a question concerning general usability and functionality of the keyboards for the official writing system. This question concerned all electronic devices instead of just the mobile phones. With this question I wanted to know if the electronic devices have provided sufficient and easy means for the writers of non-Latin alphabet to write and express themselves. The question, however, was very general which means that people might have interpreted it in various ways.



*Figure 5. The general functionality of keyboards for the official writing system of the language of the respondent on electronic devices.*

Out of the 141 who answered to this question, 50 people say that the keyboard for their language’s writing system works very well and 51 say that it works pretty well. This means that 101 people and over two thirds of the respondents are quite satisfied with the functionality of the keyboards in general. Considerably less people think the keyboards work quite poorly (13) or very badly (3). All groups except the “Thai” group have at least one

respondent who has replied “quite poorly”. The ones who answered “very badly” are one Farsi speaker, one Serbian speaker and one Russian speaker. If these three were speakers of a same language, some careful speculations could have been made but them being representatives of different languages such speculation is not possible. 11 out of the 141 respondents say that the functionality varies depending on the situation and 13 say that they do not know what to answer. The five Farsi speakers all had a different answer to this question which emphasises the individual and subjective nature of the questions.

Out of the 139 who answered to a question about the availability of different keyboards, 85 feel that the Latin script is more easily available than the keyboard that would be needed to write the respondent’s language. Therefore, according to 54 people, which is less than a half of the answerers of this question, the non-Latin script keyboard is more easily available. Among the Russian, Farsi and Taiwanese Mandarin Chinese speakers the majority has answered “no” to this question. Among the rest of the languages groups the majority has chosen “yes”.

## **4.2 The Interviews**

The four interviews were made in order to understand the perspective of an individual more profoundly. The interviewees did the online questionnaire like all the other respondents. I then collected their answers and asked some additional questions related to the theme. The aim was to have a laid-back and semi-structured interview for gathering personal opinions and attitudes towards the use of different writing systems.

### **4.2.1 Russian (Cyrillic)**

The Russian speaking interviewee is 24 years old and she lives in Russia. She is a university graduate who works as a journalist. Her answer for whether she uses the Latin script when writing Russian was “no”. During the interview, she however admitted that she has sometimes used the Latin script when teaching Russian to some foreign friends. She also has used the Latin script when writing her personal information when travelling abroad. If the Latin script was needed to make the message understood to the reader, she would use it.

When I asked the interviewee if she had heard about the Latin alphabet use while texting just because it is cheaper, she told an interesting story. When she was younger, her father used to pay for her phone bill, and he said that she should use the Latin alphabet when texting

because it was cheaper. However, she only wrote with the Latin alphabet to her father so that he would not realise that she used the Cyrillic script when writing to everybody else.

The Russian interviewee understands that when she goes abroad, the Cyrillic script might not be available on computer keyboards. In such situation she has to change the keyboard layout. The extra effort needed for the possibility to write in the Cyrillic script might feel annoying to her. Even after the extra effort, she might press the wrong key because the physical keyboard does not have Cyrillic labels. In these cases, she would probably choose to write in the Latin script since it would be less troublesome.

According to the Russian speaker, using the Latin script is not cool but she mentions that some years ago in Russia it was popular among bands to have their names in the Latin script. Also, some companies in Russia have their name in the Latin script and they do not allow the media to transliterate the name into Cyrillic script. In Russia the Latin script can be seen on street signs in places where there are tourists, like in Moscow and St. Petersburg. Some brands like Coca-Cola are written in the Latin script. Some international restaurant signs like McDonald's is written both in the Cyrillic and Latin scripts.

According to the interviewee, people in Russia do not really have the need to write in the Latin script since the Cyrillic script is understandably more easily available there. It is also considered to be the best writing system to write Russian. Using the Latin script mainly happens out of laziness to switch the keyboard into the Cyrillic script. The Latin script does not have a reputation of a “cool” way to express oneself.

#### **4.2.2 Farsi (Arabic)**

The Farsi speaking interviewee is a 33-year-old university graduate, who works in IT and data communications. He currently lives in Finland. The interviewee uses the Latin script at least once a week when chatting with friends, writing to foreign friends and when transliterating Farsi to someone who is learning it. Mainly the one to whom he writes in the Latin alphabet is his Finnish spouse and some Iranian friends who also live in Finland. For example, when chatting with friends some Finnish street names need to be written in the Latin script.

The interviewee started using the Latin script when he was a teenager and began visiting chat groups online. Back then, keyboards did not support the Farsi Arabic script, so he had to use the Latin script. When the Farsi script became available, people did not use it at first because

they used a lot of slang words which were easier to write in the Latin alphabet. Also, nowadays the interviewee feels that sometimes it is easier to write to friends in the Latin alphabet. Since the Farsi Arabic script is written from right to left and the Latin from left to right, code-switching can be tricky because the writing device can become confused about the direction changes. According to the interviewee a word written in the Latin alphabet sometimes goes automatically to the beginning of the sentence even though it is supposed to stay in the middle. This understandably can be an annoying issue.

When he writes Farsi to his spouse, who is still learning Farsi language, he sometimes writes it in the Latin script so that his spouse would see how the words are pronounced. Outside the countries where the Arabic script is known, the interviewee uses the Latin script for writing his information in hotels etc. He chooses the script according to who is going to read the text.

Before Internet-connected texting applications, Farsi speakers used the Latin script when writing SMS messages since it was cheaper to text that way. Nowadays, however, texting by using Internet-connected phone applications is unlimited so people do not use the Latin script because of this reason anymore. According to the interviewee, some younger people use the Latin script because they want to feel closer to the western countries. In the interviewee's opinion this kind of thinking is stupid, and he prefers to use the Farsi script in social media in order to encourage others to use it in writing.

In Iran on street signs, some things are written in the Latin script. They are loanwords from English and not romanised Farsi words.

The functionality of the Farsi keyboard depends on the device with which Farsi is written. Mainly, writing on computers works well but, on some phones, writing is harder since they do not support the Farsi language very well. The additional Farsi letters to the Arabic script require the Farsi writer to use key combinations. Word prediction is available for the Arabic language but not for Farsi. According to the interviewee writing in Farsi on electronic devices is not difficult but it is not easy either.

#### **4.2.3 Cantonese (traditional Chinese characters)**

The Cantonese Chinese interviewee is a 44-year-old university graduate who works in customer service. She is originally from Hong Kong but now lives in Finland. The interviewee sometimes uses the Latin script when writing Cantonese. However, with family

members she does not use the Latin script, and she prefers to use the traditional Chinese characters when communicating for example with her mother. She mainly uses the Latin script when writing with her friends. In Hong Kong almost everybody can speak English so code-switching when chatting is very common. The interviewee uses the Latin script to write some filler words of the Cantonese Language, like “*wor*, very good” or “he got a job offer *la*”. These are used to make the sentence sound more Chinese and to emphasize the meaning.

Some of the friends of the interviewee do not know how to use the traditional Chinese character input system very well so they prefer to write in English. This does not mean that they would not know how to speak Cantonese, the problem is only in the complexity of the input system on the Cangjie keyboard. Also, the interviewee’s Hong Kongese husband, who has lived in Finland for 20 years, does not know how to use the input system. He can read the characters but prefers voice messages or writing in English. According to the interviewee who can use the input system, writing in the Latin alphabet still feels easier. Since the input system is based on the smaller radicals and their combinations to create a character, the interviewee needs to use a lot of thinking in order to visualise the character and its components.<sup>4</sup> If the interviewee is lazy or in a rush, she often writes in English.

She had not heard of people choosing to write in the Latin script when texting because it is cheaper.

For the use of the Latin script in Cantonese language, people copy expressions from each other and use them in situations where they think that the reader would understand them. Cantonese has no standardised spelling rules when it comes to writing it in the Latin script. Probably because of this, the number of tones, the difficulty to write tones and good knowledge of English result to the fact that the Cantonese speakers do not write their language often in the Latin script but choose between writing Cantonese with characters and writing English.

Writing Cantonese on the phone or other electronic devices works quite well, even though it might take some time, according to the interviewee. She learned the input system when she was about 20 years old, but some of her friends never learned it correctly so they prefer to use English. The interviewee tells however, that young Hong Kongese people nowadays prefer to

---

<sup>4</sup> See more information about the character input of the Cangjie keyboard in Figure 1 and its information in Subchapter 2.2.4.



use Cantonese more than the older people. Younger people possibly do not feel that the input system is too complex for using it frequently.

#### **4.2.4 Thai (Thai)**

The interviewee of Thai language is a university graduate and she works in human resources and recruitment. She is 30 years old and lives in Thailand. She does not use the Latin script when writing Thai. The only exception for this is when she needs to write something to a person, who does not know the Thai language or its letters. Another reason to use the Latin script would be the need to write personal information while travelling abroad. According to the interviewee, writing Thai in the Latin alphabet is quite difficult because of the tone system of Thai. Therefore, it is better to write the language in Thai alphabet. When chatting with friends she uses the Thai alphabet when writing some English words.

She had not heard about the fact that people would choose to write in the Latin script when texting because it is cheaper.

In Thailand some international brands like KFC and McDonald's are written in the Latin script so people do recognise and use these Latin alphabet names. Street signs, names of places and some shop products are written both in Thai and in English, but not in Latin transliteration of Thai. In Thailand, people call Thai written in the Latin script "Karaoke language". However, there is no standardised conventions for writing Thai in the Latin script that would be in active use. If the interviewee taught Thai to someone, there would be no rules to follow how to write in the Latin alphabet. Similarly, without rules, reading this kind of text would be difficult.

The interviewee has seen Thai written in the Latin script by others. In karaoke rooms some songs have Latin alphabet transliterations for foreigners. Some friends of the interviewee, who are for example partly Thai, write certain words in the Latin script on social media because they do not know the Thai writing very well.

According to the interviewee, Thai alphabet computer keyboard works very well but phone keyboards could be better. She says that on a phone it is difficult to switch the keyboard and to remember where all the keys are located. Nonetheless, she is used to typing so she does not have many problems when it comes to writing Thai on the phone.

### **4.3 To conclude on results**

In the previous subchapters of Chapter 4, I introduced the quantitative data selectively, since it turned out that not all the questions of the online questionnaire are relevant to this study. However, now all the most relevant results of the quantitative data are introduced. In addition, the four interviews are presented in detail. The open question answers of the questionnaire will be discussed in the following analysis Chapter 5.

## 5 Analysis

In this chapter I look into the main findings of the data selectively in more detail. I compare the quantitative data, the open question answers and the interviews to each other.

Additionally, results of previous studies related to the topic are discussed. The focus is on the separate groups that have been introduced earlier: “Cyrillic”, “Arabic”, “Chinese”, “Thai” and “Others”. The “Cyrillic”, “Chinese” and “Others” groups are further sectioned into smaller groups by language or by geographical speaking areas of the languages. First, I discuss the analysis from the point of view of these five groups and in the end of this chapter (5.6) I introduce some findings that concern more than one group.

Throughout this chapter I use direct quotes from the respondents to emphasise the phenomena and to bring the respondents closer to the reader. The quotes support the findings that are also visible in the quantitative data and in the interviews. All the quotes are represented in their original form, without spelling or grammar corrections. Some clarifying additional notions are marked in square brackets []. The quotes are numbered, and I refer to them in the text by the number.

### 5.1 Cyrillic and Latin scripts

The “Cyrillic” group stands out from the other groups by the amount of people who use the Latin script. This is especially apparent among the speakers of Serbian, Serbo-Croatian and Bulgarian, as clear majority of them say that they use the Latin script when writing their language. The Russian speakers, however, are an exception to this by having a small majority who do not use the Latin script when writing Russian.

#### 5.1.1 Russian, Bulgarian and Chechen

Here I analyse the data gathered from Russian, Bulgarian and Chechen speakers. They all employ the Cyrillic writing system which can result in some similarities in the answers among the speakers of these languages.

Around half of the respondents of these languages use the Latin script themselves and all of them have seen it used by others. Writing information about oneself when travelling abroad is the most popular domain to use the Latin script among these respondents. The most popular

reason for using the Latin script is the unavailability of the Cyrillic keyboard and the need for non-native speakers to understand writing for example in teaching situations.

The Russian interviewee has similar thoughts about the use of the Latin script as the majority of the other respondents in this group. Speakers of Russian, Bulgarian and Chechen use the Latin script in limited ways. The Latin script is used abroad for the readers to understand the written text, but in the countries where these languages are spoken and written, it is not necessary to use the Latin script. The Latin script is not considered superior to the Cyrillic script and the Cyrillic script is mostly thought to suit the languages in question better than the Latin script.

The availability of the Cyrillic script keyboard however, divides opinions. The interviewee says that the Latin script keyboard is more readily available on electronic devices and so says 17 other respondents in the questionnaire. 12 say that the Latin script keyboard is not more easily available than the Cyrillic keyboard, as said in quote 1.

- (1) Bulgarian speaker, woman, 38, Ireland: *“I wish cyrillic was more easily accessible on all electronic devices.”*<sup>5</sup>

Some people think that the Latin script is not pleasant to use (quote 2) or that it gives the reader an uneducated impression of the writer (quotes 3–4).

- (2) Russian speaker, woman, 48, USA: *“transliteration is not pleasant to use.”*
- (3) Russian speaker, man, 33, Turkey (Kazakhstani): *“except for trademarks and product names it is considered bad manner to use latin alphabet”*
- (4) Bulgarian speaker, woman, 33, Bulgaria: *“It is considered careless/uneducated to use the Latin alphabet even on Facebook.”*

### 5.1.2 Serbia with digraphia

In this subchapter I analyse the data given by 33 Serbian speakers and two Serbo-Croatian speakers. These two language variants are discussed in the same group, since the two Serbo-Croatian speakers both use the Latin script actively alongside with the Cyrillic script. This indicates that the Serbo-Croatian speakers are influence by the digraphic situation that exists in Serbia.

---

<sup>5</sup> This is the first of the several quotes that are presented in this study. They are direct quotes of open question answers, written by the online questionnaire respondents. The sentences are not corrected or edited.

In Serbian both the Cyrillic and Latin alphabet are used in various domains, which makes Serbian language a playground for actively working digraphia (Ivković 2013: 335). In the data there are different kinds of answers also among the Serbian speakers, but the tendency is that the Serbians do use the Latin script regularly and in various domains, as noted also by Ivković (2013: 335). Out of the 33 Serbian and two Serbo-Croatians speakers, four respondents say that they do not use the Latin alphabet when writing their language.

Macedonian language has almost identical romanisation conventions with Serbian with which Macedonian shares the usage of the Latin script. There is a similar kind of situation between Bulgarian and Russian which share a lot of online romanisation patterns together. Ivković's suggestion is that these two sets of Slavic languages are in the opposite sides of a continuum: Serbian and Macedonian are heavily romanised whereas Bulgarian and Russian are much less so. (Ivković 2015)

The data gathered for this research partly indicates the same results. Out of the Russian speaking respondents, 40.0% use the Latin alphabet when writing their language. A clearly larger number of Serbian and Serbo-Croatian speakers (88.6%), use the Latin alphabet. However, in contrast to Ivković's statement, the data of this study show that out of all the Bulgarian speakers who participated in this study, 87.5% use the Latin script when writing Bulgarian. Ivković's theory does not entirely apply to this set of data, since the use of the Latin script is so high among the Bulgarian speakers.

As stated earlier in the text the Cyrillic script and the Latin script have different nature and functionality in the eyes of the Serbian and Serbo-Croatian writers. Many people write in the Latin script because it is a habit, but the Latin script is also found to be more neutral than the Cyrillic script. The Latin script seems to be somewhat preferred over the Cyrillic script in some situations because the Cyrillic script is mostly used by the Orthodox Church (quote 5) and using it might also be considered political (quotes 6 and 7).

(5) Serbian speaker, woman, 30, Switzerland: *"Latin script is very much used in Serbia, and it is mostly the Orthodox church that tends to avoid it."*

(6) Serbian speaker, woman, 26, UK: *"Using cyrilics is considered a statement of pride and nationality. Although I am proud of my country and heritage I do not wish to be seen as a nationalist."*

- (7) Serbo-Croatian speaker, woman, 35, Finland: *“it is common to do so [to write in the Latin script], if you use Cyrillic on social media, that normally reveals conservative political views”*

Some respondents however prefer the Cyrillic script when writing by hand (quote 8) or when writing in the cursive by hand (quote 9). One person is worried about the disappearance of the Cyrillic script because of the dominance of the Latin script (quote 10). One Serbian respondent says that she prefers to use the Cyrillic script because she already uses the Latin script for writing the other languages that she knows. Another Serbian speaker says that the Cyrillic script represents the phonology of the language better than the Latin script. For many using the Latin script is a habit (quote 11). One person chooses to write in the Latin script so that the speakers of neighbouring countries could understand the written text (quote 12).

- (8) Serbian speaker, woman, 30, Switzerland: *“When writing by hand, I never use the Latin script, only when typing as it comes easier.”*
- (9) Serbian speaker, man, 40, Finland: *“Ever since I was a child I found using Latin easier when i had to handwrite in capital or block letters. However, if I switch to cursive I find Cyrillic easier and more fluent.”*
- (10) Serbian speaker, woman, 21, Serbia: *“I have nothing against Latin letters, but they will delete Cyrillic at this rate. I feel like more people are using Latin over Cyrillic which makes me sad.”*
- (11) Serbian speaker, man, 46, Serbia: *“it is a habit from elementary school. my mat teacher ask me to write the latin.”*
- (12) Serbian speaker, man, 48, Serbia: *“In Serbia we use officialy both letters, so for most of us it is the same. But people from the region (Croatia, Bosnia) understand language, but write latin, so it is easier to understand each other.”*

The Serbian and Serbo-Croatian speakers like writing in the Cyrillic script more than in the Latin script. 94.3% of the respondents like writing in the Cyrillic script but only 65.6% like to write in the Latin script. It seems that the Serbian and Serbo-Croatian speakers use both scripts quite equally, but some people have personal preferences to use one over the other.

In the quantitative data the Serbian respondents stand out as the biggest group of speakers of one language. The Serbian speakers stand out also because the individuals in the group have chosen more multiple-choice answers than the speakers of other languages. They use the

Latin script a lot more and, in more domains, than respondents who represent other languages than Serbian or Serbo-Croatian. According to the respondents' questionnaire answers, the Latin and Cyrillic script in Serbian are often interchangeable. The style of writing in the Latin script can be formal or vernacular. Thus, the digraphic situation in the Serbian language community is well reflected in Serbian speakers' answers.

Two Serbian speakers say that they use the Latin script when sending SMS messages because texting in the Latin script is cheaper than texting in the Cyrillic script (quote 13). Because of the length of the binary codes of the graphemes (as explained in Subchapter 2.2.2) one can write longer messages in the Latin script compared to other scripts. Some respondents of other languages mentioned this reason too, but in the case of Serbia, where the Latin script is equally usable with the Cyrillic script, the writer might be even more motivated to write text messages in the Latin script.

(13) Serbian speaker, woman, 33, Serbia: [Uses the Latin keyboard] *“because they charge me more if I use cyrillic letters in sms. Stupid, i know.”*

## 5.2 Arabic and Latin scripts

The “Arabic” group of this research is a heterogeneous group that consists of Farsi and seven distinct dialects of Arabic language. However, they all utilise the Arabic script as the official writing system and therefore these languages share common challenges and sociolinguistic settings. Over a half of the respondents in this group use the Latin script themselves when writing their language and all of them have seen it used by others.

In Palfreyman & Al Khalil (2007) university students in the United Arab Emirates mention several reasons why they would use the “English Arabic”: it is easy and fun to write, and the Latin alphabet has positive social connotations since only educated people use it. A surprisingly large amount of young United Arab Emirate people use the Latin alphabet to write secret messages to one another at school, since the teachers of older age do not understand the romanised Arabic texts. (Palfreyman & Al Khalil 2007: 59–60)

When comparing the heterogeneous “Arabic” group and the young students of the United Arab Emirates in Palfreyman & Al Khalil (2007), it can be stated that in the case of the Arabic script, the reasons to use the Latin script differ from language community to another. In contrast with the students in the United Arab Emirates, only a few respondents of this

study say that they use the Latin script because it is easier or faster. The most common reason to use the Latin script is the unavailability of the Arabic keyboard. Some do mention that using the Latin script does not demand using spelling rules, which makes it more carefree to use than the Arabic script. None of the respondents mention that the Latin script could be used as a secret written language. However, during the interview, the interviewee admitted that Farsi messages written in the Latin script written would work as a secret text that would be unreadable for the old people who do not know the Latin alphabet.

In the Palfreyman & Al Khalil research (2007) United Arab Emirates' teenagers had a very positive and social view on the use of the Latin script. The respondents of this research, however, have a more neutral or even slightly negative approach to the use of the Latin script. Some respondents prefer to use it when chatting (quotes 14 and 15) and for some the use of the Latin script on electronic devices is just a habit from the times when the Arabic keyboards were not available (quote 16). One respondent mentions that the Latin script is not very suitable for writing Arabic and it is often considered a threat to the Arabic written language (quote 17).

(14) Farsi speaker, man, 32, Finland: *"In social media, sometimes Iranian use english writing system, penglish."*

(15) Moroccan Arabic speaker, man, 21, Morocco: *"in chatting, using latin alphabets is more easy than arabic."*

(16) Farsi speaker, man, 31, Finland: *"Old habits die hard. At early days of home computers and cell phones the Persian characters were not widely available in most sw products (or were not sophisticated). So we were restricted to innovate."*

(17) Algerian Arabic speaker, man, 25, Algeria: *"It does not always read well. Besides, Muslim-Aravs believe that such practices are likely to undermine Arabic (the language of Quran). Some consider that this is a sign of the degradation of the Arabic language."*

Even though the respondents acknowledge that the Arabic keyboards have developed and that the writability of languages that use the Arabic script has increased, only five respondents of this group think that the Arabic script is more easily available than the Latin script. These people think that the Arabic script is accessible, and they seem to be satisfied with the situation (quotes 18 and 19).



(18) Jordanian Arabic speaker, woman, 19, Jordan: *“The Arabic keyboard can be installed on any smartphone or tablet, and laptops sold in the Arab world come with both Arabic and English stickers on the keys. Both the Latin and Arabic alphabet are easily available.”*

(19) Farsi speaker, man, 28, Germany: *“Latin used to be more easily available but nowadays access to Persian keyboard is possible on pretty much all devices.”*

However, some respondents are still not pleased with the functionality of the Arabic keyboard when writing their language (quotes 20 and 21). Also, the Farsi speaking interviewee thinks that there is still room for further development of keyboard technology. When writing with two scripts on the phone the words might switch places against writer’s intention and word prediction is still not optimised for Farsi language.

(20) Farsi speaker, man, 31, Finland: *“Although modern tech widely support Persian keyboard, but it's not always about availability. Note that things like direction of writing and how letters get combined into words can make handling things very complicated. For example, Persian (and Arabic) are written from right to left while numbers are written from left to right. Many SW products even these days fail to correctly handle a Persian sentence that is written with Perso-Arabic system and contains some numbers and Latin based words.”*

(21) Standard Arabic speaker, man, 21, USA: *“Unfortunately with Arabic, and its extended characters, I find it extremely hard to find support for them, especially when writing in dialect or for a smooth pronunciation of foreign names, ideas or concepts. And latin-based script is difficult to read or understand because of sheer lack of extensive characters that are semi equivalent to Arabic letters. We often, if need be, result to numbers and punctuation for approximation”*

In the Arabic culture the vernacular languages have had an inferior position in the society, but the prestige status of English elevates the status of these vernacular languages through the use of the Latin script when writing ASCII-ised Arabic (Palfreyman & Al Khalil 2007: 61).

Therefore, also the vernacular language variants are accepted to be used if they are written in the Latin alphabet. The respondents of this research almost exclusively use the spoken language when writing with the Latin alphabet which is connected to the findings of Palfreyman & Al Khalil (2007). In addition, the majority of the “Arabic” group says that they use English words if they write their language in the Latin script. This indicates that the

influence of English on these languages is real even though the questionnaire does not give answers to whether English has a prestige status in the language communities or not.

Some respondents say that when they were younger, they thought that using the Latin script was cool but today the opinions have changed to neutral or negative (quote 22). This indicates that the opinions on writing systems are changing and they might be affected by the global atmosphere online.

(22) Jordanian Arabic speaker, woman, 19, Jordan: *“Arabic is written everywhere in arabic letters, except on social media, people started to write with latin letters, i don't why, but when i was younger i thought it was cool but now i find it so ugly and only use it with my friends who write arabic in latin letters (some of them are not used to the arabic keyboard)”*

### **5.3 Chinese and Latin scripts**

The Chinese characters are used quite differently between different languages. For example, in mainland China the Mandarin Chinese utilises the simplified characters whereas in Taiwan, Mandarin Chinese and other language variants spoken on the island are written with the traditional Chinese characters. The Cantonese Chinese language also utilises the traditional characters. Japanese uses the Chinese characters too, alongside with the Hiragana and Katakana syllabic scripts but as mentioned earlier in the text, Japanese is in the “Others” group (see Subchapter 5.5.1 for the analysis concerning Japanese).

#### **5.3.1 Simplified characters (Mainland China)**

For this research, nine Mandarin Chinese speakers who use the simplified characters gave their answers. Out of these nine, five people live in mainland China and the other four live in different countries. Therefore, this subgroup consists of respondents from five countries but here I concentrate on the situation in mainland China.

Majority of the respondents using the simplified characters use the Latin script themselves. Almost as many respondents, with only one addition, have seen the Latin alphabet used by others for writing Mandarin Chinese. The respondents mention the use of Pinyin several times. They use Pinyin especially in association with typing Chinese on electronic devices or when transcription of the language is needed. Some of the respondents who mention the

Pinyin, say that it is the most convenient and the fastest means for writing Chinese on electronic devices.

When it comes to the linguistic situation in mainland China, the Latin alphabet and the Chinese characters have a special relationship. The phonetic Pinyin system is widely used alongside the Chinese simplified characters. In Chinese, some words exist only in the Latin alphabet for example “KTV” (karaoke). Even though some Chinese people might not know English, they do know the letters that English utilises since they are the same letters as those in Pinyin. Therefore, English and Chinese partly share the same script but not the same orthography or pronunciation. To some extent, it could be even said that Chinese has digraphia. (Hansell 2015)

Due to the strict Internet censorship in China, certain words connected to unpleasant events or criticism against the state are blocked on the Internet. To circumvent censorship, people need to find new ways to express these banned matters online by creating new character combinations or by using the Latin alphabet. New invented word forms and conventions fool automatic systems that search for forbidden words online. For example, when the word “government” (政府 zhèngfǔ) got banned, people started to use the Pinyin initials “ZF” to refer to the government so that the conversations online could go on. (Qiu 2008)

One respondent mentions this censorship in the open question answers (quote 23). Using the Latin script to create new ways to write words is innovative and an interesting phenomenon. Mainland China is perhaps a unique example of these kinds of censorship-driven scriptal innovations.

(23) Mandarin Chinese speaker, man, 47, China: *“a way to input Chinese characters or to express something forbidden on purpose.”*

### **5.3.2 Traditional characters (Taiwan)**

In this chapter I look into the Mandarin Chinese spoken in Taiwan, which utilises the traditional Chinese characters. Eight speakers of Taiwanese Mandarin Chinese participated in this research. Four of them used the Latin script to write Chinese and five have seen it used by others.

The phonetic typing system Zhuyin Fuhao or “bopomofo” that is used in Taiwan, is not an independent writing system, but a tool for learning the Chinese characters and to type them on electronic devices (Su 2001: 110). Zhuyin Fuhao is based on the phonetic symbols of Chinese characters but not on the Latin letters as is the case with Pinyin. Even though Pinyin is to some extent used in Taiwan, according to at least two of my respondents, the Latin keyboard and the Pinyin writing system are not the preferred means for writing Mandarin Chinese (quotes 24 and 25). The respondents are not used to typing Mandarin Chinese in the Latin script and they feel that it is too rigid.

(24) Mandarin Chinese speaker, woman, 23, Taiwan: *“Just not so familiar with Latin keyboards.”*

(25) Mandarin Chinese speaker, woman, 22, Taiwan: *“To me, it is the most familiar system, and it's easier to use. It often takes a lot of time if using Pinyin.”*

According to the research of Su (2007: 80), in Taiwan, typing English is perceived to be easier than inputting Chinese characters, which affects the choice of script. In this research, only two Taiwanese respondents say that they use the Latin alphabet because it is faster and easier. Two respondents say they use the Latin script because of the shorter binary code and two use it because some things are not translated to their language. None of the Taiwanese respondents give extra information in the open question answers whether they would prefer the Latin script for other reasons. Because of the scarcity of the replies for the multiple-choice questions and the fact that the only open question answers by Taiwanese respondents tell about the preferences to use the non-Latin script, this research indicates opposite results to those that are found in Su (2007: 80). However, more information about the possible role of English in the respondents’ Latin script writing would be needed to make final conclusions on this matter.

### **5.3.3 Traditional characters (Hong Kong)**

In this chapter I focus on the Cantonese Chinese that utilises the same traditional Chinese characters as the Mandarin Chinese spoken in Taiwan. Four out of ten Cantonese Chinese speakers inform Hong Kongese as their nationality, other five being Chinese. Five out of the ten Cantonese speakers use the Latin script when writing Cantonese and six have seen it used by others.

Even though 90% of the population in Hong Kong speaks Cantonese Chinese, English is still considered a very important language of use. This is not a surprise since English has been an official language in the area since the early 1970's due to the British colonial regime. The bilingual history is apparent in online communication as many Hong Kongese prefer to mix both Chinese and English as opposed to writing with just one language. (Lee 2007: 186, 188)

All of the six respondents who say that they have seen the Latin script used by others inform that they have seen English loanwords written in the Latin script. In addition, the Cantonese speaking interviewee mentioned several times the use of English in the instant messaging. Therefore, the findings of Lee (2007: 186, 188) are in accordance with the findings of this study.

The romanised writing for Cantonese is not fully standardised (Lee 2007: 192) which affects the usability of the Latin script for Cantonese. The interviewee as well as one online questionnaire respondent say that using the Latin script for Cantonese is tricky since there are no standard spelling rules. While writing in the Latin script causes some trouble, the Hong Kongese writers encounter typing complexity and tardiness when using the character input systems as well (Lee 2007: 194). The Cantonese interviewee mentions the same problem when it comes to typing Cantonese on a phone. Typing of the traditional Chinese characters on the Cangjie keyboard demands thought and a good visual memory. In some cases, it is easier to code-switch to English than continue writing in the characters. When the shape of a character is forgotten, Cantonese words written in the Latin alphabet are preferred, as mentioned by one respondent:

(26) Cantonese speaker, woman, 33, Hong Kong: *“Some people may write some characters in alphabets in a sentence, maybe because it is easier to write or forget how to write the character. In this case others still get what they mean because they know the context in the conversation/text”*

#### **5.4 Thai and Latin scripts**

Three out of six Thai respondents use the Latin script when writing their language and all of the six respondents have seen it used by someone else.

In Tagg and Seargeant (2012) it is noted that young Thai-English bilinguals choose to be playful, not only with English but also with the “Karaoke language”, a romanised Thai. As

young international Thais they share an identity that combines the two languages as well as the two scripts. They use the Latin script, orthographic manipulation and unconventional writing styles to express playfulness, but also more serious things can be expressed with these techniques. (Tagg & Seargeant 2012: 210–212)

One could assume that monolingual Thai speakers would use the Latin alphabet for writing Thai less than those who regularly write in English. The respondents of this research all know good English and monolingual Thai speakers have not participated in this study so this hypothesis cannot be tested.

Since the Thai speaking interviewee is an internationally active person who uses English every day at work, she is very familiar with the Latin alphabet. However, she feels that the Thai alphabet suit the Thai written language much better and she sees only few reasons to use the Latin alphabet for writing Thai. On the other hand, in the questionnaire, those who use the Latin alphabet, chose quite many domains of use for the Latin script. The respondents also have seen the Latin script used in relatively many contexts.

In Thai language the tones play a fundamental role. The Thai script has tone marking integrated into the syllables and letters. If the Thai language was written in the Latin script, the tones should be written separately as diacritics or in some other ways. One respondent pointed this out in the online questionnaire (quote 27) and the interviewee also discussed about this disadvantage of the Latin script when it comes to writing Thai.

(27) Thai speaker, woman, 30, Thailand: *“In Thai, tones distinguish the meaning. It is impossible to clarify meaning without tone marks.”*

The Thai speakers seem to be very satisfied with the Thai script and how it is used in the language. The interviewee mentioned that typing Thai on phone keyboards can be occasionally time consuming but if the writer becomes used to the keyboard typing should not cause any bigger problems.

## **5.5. Others and Latin scripts**

In the previous chapters the four previous groups concerned scripts that the four interviewees use for writing their languages. The group “Others”, that is discussed in this chapter, contains scripts that are not represented by an interviewee. “Others” group consists of several different

languages and language scripts. Japanese and Hindi are discussed in their own subgroups because they are represented by larger amount of questionnaire respondents than other languages in this group.

### 5.5.1 Japanese

Three out of six Japanese respondents use the Latin script themselves and all of them have seen it used by others when writing Japanese.

In Japanese texts, English loanwords such as *snack* and *make* are mostly written in their English form rather than in the Japanese romanised orthography *sunakku* and *mēku* (Backhaus 2006: 144). In the questionnaire there were no direct questions regarding this kind of Latin script use, but one respondent mentions that if he writes loanwords in the Latin script, he would generally write them in the original spelling form. Therefore, it is possible that the other Japanese speaking respondents use the Latin script in the same way. Another matter concerning loanwords is that the respondents have seen Japanese words written in the Latin letters in words that have been borrowed by English language (quote 28).

(28) Japanese speaker, woman, 24, Finland: *“I see some Japanese word which is used as English (sushi, tsunami, umami, ramen, miso, karate...)”*

For a non-native Japanese speaker using the Latin script is a way for easy and fast way to write Japanese (quote 29). For language learners of Japanese, the complex three-script-system of the Japanese orthography can seem troublesome to learn but the romanisation of Japanese has enabled second language learners to acquire the language without too much effort. One Japanese native speaker has seen some second language learners use the Romaji for everything in online communication (quote 30).

(29) Non-native Japanese speaker, man, 70, USA: *“I also speak some Japanese and \*always\* use the Latin, because I'm too lazy to learn three non-latin alphabets for a language rapidly adopting Latin.”*

(30) Japanese speaker, man, 33, Finland: *“While officially only foreign words should be written in romaji, online you can find foreigners attempting to use it for everything. Ironically I myself am guilty of this at the present as "romaji" itself is a bastardization of "Roman signs".”*

### 5.5.2 Hindi

Four out of the six Hindi respondents use the Latin script when writing Hindi and all of them have seen it used by others. All the Hindi speakers use informal language if they write Hindi in the Latin alphabet.

The influence of English language in Internet communication has affected the Hindi speakers notably. The Devanagari keyboard has not always been easily available which means that some Hindi speakers have learned to write Hindi in the Latin script and continued to use it even after the Devanagari keyboards became more available (quote 31). For one questionnaire respondent, English has been the medium of education which has led her into a situation where now she feels that reading and writing Hindi in the Latin script is easier than in the Devanagari script (quote 32).

(31) Hindi speaker, man, 23, India: *“When I started texting, keyboards weren’t (easily) available for Devanagari, and when they became available, I had no practise with typing in Devanagari and it was convenient to just write in the Latin script.”*

(32) Hindi speaker, woman, 25, Finland: *“The medium of education was English throughout my studies. So naturally reading and writing Latin script was easier than Hindi (Devanagari Script) for me.”*

### 5.5.3 Miscellaneous

In this subchapter I look into the languages and scripts that are represented by few people in the data. Each of the 13 languages that are considered in this subchapter are represented by only one or two respondents. Therefore, it is not possible to make generalisations about the use of the Latin script in these languages.

One answer by a Georgian respondent is related to the understandability of text (quote 33). The respondent’s decision to use the Latin script seems to be inspired by the fact that social media platforms have instant translation options for posts, captions and other texts alike. If text is written in a “wrong” script, translating machines are unable to translate the text correctly or fail to do it at all. Therefore, the Latin script can be used to filter out unwanted readers who do not understand the text without the machine translations.



- (33) Georgian speaker, woman, 32, Georgia: *“Sometimes, when I don't want some of my non-Georgian friends to understand what I am talking about I post a status written in Latin alphabet :))”*

One respondent pointed out an interesting reason to use the Latin script when writing her native language, Bengali. She stated, that when she writes Bengali on a phone, she feels that the language should be informal spoken language and writing the spoken language in Bengali script would feel weird. Therefore, she prefers to write quick informal messages in the Latin script (quote 34).

- (34) Bengali speaker, woman, 23, USA: *“Nowadays it feels though that if I go to the trouble of writing a text in Bengali script to my family, I might as well use the standard dialect - but sometimes it feels a little unnatural to do that so I use my spoken dialect in Latin script. I guess I could write it with Bengali script? But somehow it feels weirder than doing it in Latin script.”*

Some of the languages presented in the data do not have sufficiently well-functioning keyboards available. For example, respondents speaking languages like Nepali, Tamazight, Javanese and Marathi use the Latin keyboard because the ones designed for their languages are either not easily available or they are difficult to use (quotes 35–37).

- (35) Javanese speaker, man, 23, Indonesia: *“Support for Javanese keyboard is almost non-existent. Most people prefer Latin because it is easier. Javanese is only recently included in Unicode, so widespread support is currently still minimal.”*
- (36) Nepalese speaker, man, 35, Finland: *“I don't use Devnagiri at all on my phone and i don't know...if such service is available.”*
- (37) Marathi speaker, man, 30, India: *“English typing is much easier than typing Devanagari script.”*

## 5.6 Comparing and contrasting the scripts

In this chapter I go through some results that are not restricted to any specific language or language group. I compare and contrast different languages and their scripts, in order to find explanations to the differences and similarities between them.

In the quantitative data the contrast between the Latin script writing styles of the Japanese speakers and the speakers of languages spoken in India and Bangladesh is quite interesting. While all the Japanese speakers use an official spelling convention and use both standard and spoken language, the speakers of Indian languages do not follow the standard (except for one who does not know) and they use only spoken language. These situations might spring from the fact that in Japanese, the Romaji writing convention is well known, whereas in the languages spoken in India and Bangladesh, official broadly spread romanisation conventions do not exist.

Only in the “Cyrillic” and “Chinese” groups were there respondents who used the Latin script exclusively when writing formal standard language. In the other groups the respondents use only the spoken language or both spoken and formal language. From these results it can be assumed the languages that utilise the Cyrillic or Chinese scripts, have possibly adapted the Latin alphabet use more extensively than those languages in the other groups.

Within the “Arabic” and “Thai” groups the answers were in some places the most unanimous. In both groups the respondents made the same choices several times with other respondents in the same group in the multiple-choice questions. The same kind of consistency within the other groups was not visible. The “Arabic” and “Thai” groups, however, were also the smallest groups. The high chance for coincidence in the results due to the small number of respondents might explain the similarity in the answers.

In the online questionnaire most of the speakers of Russian, Farsi and Taiwanese Mandarin Chinese speakers say that the Latin script is not more easily available than the non-Latin alphabet. This might indicate that in these language communities, the official keyboard is easily available on phones and computers, or at least the unavailability of the official script keyboard is not the reason for choosing the Latin script.

Russian, Ukrainian, Japanese, Cantonese Chinese and Taiwan Mandarin Chinese speakers use the Latin script less in online settings than the speakers of other languages. In the cases of Russian and Ukrainian I would assume that the Cyrillic keyboards are easily available and functional which removes the need to use the Latin script. In the language communities of Japanese, Cantonese Chinese and Taiwan Mandarin Chinese the frequent use of English, especially online, might motivate the respondents to switch straight to writing in English instead of writing their language in the Latin script. Since the respondents have not explicitly

specified why they do not use the Latin script that much when chatting online, these assumptions are entirely my own.

16 out of the 40, who answered that they write non-English foreign loanwords in the Latin script, live in a country where their language is not the main language of use. From these numbers it cannot be said that the country of residence affects the use of foreign loanwords other than English.

Two respondents, one Mandarin Chinese and one Cantonese speaker use the Latin script to write their language but have not seen the Latin script used by others. Their choice of answers is interesting since majority of the respondents have at least seen the Latin script used by others, including the respondents who do not use it themselves. These two respondents have two domains of Latin script use in common: work assignments and when translating their language. They both use the traditional Chinese characters to write their languages. The respondents have not given more specific reasons for their replies in the open question answers. Therefore, only the few commonalities between the two respondents give some explanation to why these respondents have such uncommon answers when it comes to writing their languages in the Latin script.

Nine respondents say that they neither use the Latin script themselves when writing their language, nor have they seen it used by others. These respondents are speakers of Mandarin Chinese (4), Cantonese Chinese (3) and Marathi (2). Since these respondents are representatives of just three languages it can be interpreted that in these languages the Latin script is not widely used. In the case of the Mandarin Chinese speakers, however, Pinyin is probably not considered as a separate Latin script by the respondents. In addition, if these respondents were given more examples of where they might have seen the Latin script used in their language, they might have answered differently.

According to the data, the usability of Pinyin for typing Mandarin Chinese is viewed differently in mainland China and in Taiwan. In mainland China, Pinyin is considered the fastest and the easiest means for typing Chinese whereas in Taiwan the respondents prefer to use the Zhuyin Fuhao input system since they are more familiar with it. In these language communities, the once learned typing convention seems often to be the one that is used by the respondent in the future.

The opinions on the functionality for changing keyboards on mobile phones and on the general aptitude of electronic devices to handle different writing systems must, at least partly, depend on the respondent's personal habits and needs. This stems from the inconsistency of the answers in the data between the language of the respondent, the phone brand used by the respondent and the answers concerning the use of electronic devices. In the data there were only very small amounts of similarities in answers between people, which was unexpected.

Many respondents are satisfied with the keyboards that they use on electronic devices. However, 14 people think that switching between writing systems works either “quite poorly” or “very badly” on their phones and 16 have chosen one of these two options when asked about the general functionality of keyboards on any electronic device. This finding shows that there is still some development to be done by phone designers.

The results on the general usability of keyboards on electronic devices when writing in the official non-Latin script does not correlate strongly with the results on whether the respondents like to write in the official non-Latin script. Some of the respondents who say that they do not like writing in the official script have said that the keyboards work “pretty well” or “very well”. On the other hand, some say that the keyboard works “quite poorly” when writing in the official script. Only one respondent who does not like to write in the Latin script thinks that the keyboards work “very badly”. Furthermore, relatively many respondents (11) who do like to write their language in the official non-Latin script, say that the keyboards work “quite poorly” or “very badly”. From the data of this study, it can be concluded that poor keyboard functionality on electronic devices does not affect the opinions writing in the official non-Latin script negatively. The effect of the poor keyboards is rather the opposite, since 11 out of the 16 respondents who think the keyboards on the official script work “quite poorly” or “very badly” say that they like to write in the official non-Latin script.

## 6 Discussion

In this chapter I analyse the suitability of the methods for this study and examine the expected and unexpected findings of the research. I also discuss the quality of the data.

### 6.1 On the methods

The methods were suitable for this research since they provided a good amount of data that was possible to process. The interviews and the online questionnaire were a good combination that supported each other. The data was advisable since its results answered to the research questions. The data was sufficient for this kind of master's thesis study and literature related to this topic was well available. A broader quantitative data and additional text analysis would have complemented the data, but due to the relatively dense nature of master's thesis, time and resources were limited.

The link to the online questionnaire was sent to Facebook groups and Reddit discussion boards where potential respondents could be found. Nonetheless, it was not controlled who was going to fill in the questionnaire since the questionnaire was open to everybody. Here I believe that people were truthful, and they answered the questions in earnest and not for example lie their age or native language. This trust to the respondents comes from the fact that I received only datasets with appropriate answers. None of the datasets contained pranks or other ill-mannered behaviour. In the end, the quality of data and the suitability of the respondents turned out to be more than enough for a study of this nature.

The data file that was composed straight from the server of the questionnaire (e-form by the University of Helsinki) turned out to be prone to produce duplicate answers. These duplicates made the data processing slower but did not affect the final results, since the duplicates were deleted from the data. However, human errors done especially during the analysis of the quantitative data are very possible.

The online questionnaire as well as the interviews were in English. This of course restricts the variety of respondents, which was known already before conducting the research. However, using English is still the best way to get as many respondents as possible to participate in the research. Since I myself do not know sufficiently any language that uses a non-Latin script, the means to collect data and analyse them were restricted. On the other hand, as one of the research topics of this study was online language, English, being the dominant language of

use in that domain, is quite suitable for the nature of this study. None of the respondents seemed to mind using English for filling in the questionnaire.

10 out of the total 28 questions were obligatory. The remaining 18 questions thus were optional which explains the lack of answers from some respondents to some questions. Also, the questionnaire was such that everybody could see all the questions regardless what they have answered in the questions. Therefore, even if the respondent was asked to skip some questions in case they answered “no” to a certain question, they could still accidentally answer to the questions they were supposed to skip. Consequently, there were some contradictions in the data, which eventually did not do harm the final results.

Some of the respondents may have been uncertain whether the words they write in the Latin script are actually words from their language or if they are code-switching to English. In the research, I was also interested in the code-switching part, since I interpreted that some English words here and there do not make the text written in English. However, the respondents may have interpreted that they do not use the Latin alphabet at all when they write their language, even though some code-switching occurs.

Sometimes it is impossible to determine what language is being used, since loanwords and script-switching muddle the line between languages (Backhaus 2006: 144). Some of the interviewees had to take a moment to think whether the words they have seen written in the Latin script were actually written in their language or in English. For any person not dedicated in the field, separating loanwords from a given language can be challenging.

The questionnaire had a few ambiguities to it. All of my four interviewees had to rethink some of their online questionnaire answers during the interview since they had not quite understood what was asked in the question. After some detailed explanation, they sometimes decided to change answers. Therefore, some of the online questionnaire respondents probably have misinterpreted some of the questions as well. More examples to help the respondents answer the question could have been inserted but that, on the other hand, might have led the respondents too much and compromised independence and subjectivity of the respondents' answers. In addition, because of the broadness of the research, the additional examples and guidance texts would have probably made the questionnaire too long to fill in. People answering the questionnaire took their own time to fill in the questionnaire without any

compensation. Therefore, the questionnaire needed to be compact enough for the respondents to be able to answer all the questions in reasonable amount of time.

## **6.2 On the results**

Same characteristics of Latin script use were found between previous literature and in the data of this research. However, because of the individual nature of the questionnaire answers, most irregularities can be explained by the uniqueness of each participant.

The research topic gained delightful amount of interest among people on the Internet. People are interested in these kinds of things, which probably also helped to get the relatively large number of respondents. The questionnaire collected a lot of interesting answers and ponderings from the respondents. The topic is clearly in the minds of people and more research of this kind would be welcome. Through this research I have also noticed that Internet provides excellent means to gather data. Links are easy to distribute online, and people can share them quickly with their friends. The different Internet communities support individuals in their efforts and help them to reach their goals.

It was not expected that many of the answers of the respondents were so variable. The individuality of answers can be explained by the fact that the speakers have their own preferences, their own backgrounds, different phones or social circles. All of these factors affect their choices when talking about the use of different writing systems. Therefore, by looking only at the metadata of the respondents, it is not possible to predict, who uses the Latin script and who does not. For example, age, gender or occupation does not seem to have any significant effect on whether the respondent uses the Latin script or not.

In the data, dozens of languages are represented by at least one person. Therefore, the data contains a lot of information about the script choices of individuals, each of whom have shared very interesting and useful facts from their perspective. However, the data is not big enough to make broad crosslinguistic generalisations on all writing systems. The answers of the respondents vary a lot within the same writing systems. This means that the writing system that is officially used in the language of the respondent does not always indicate whether s/he uses the Latin script or not.

On the level of different languages only the Serbian speakers were represented enough for making generalisations about their use of the Latin script. The digraphic situation in Serbia is

clearly visible in the results. Not only the Serbian speakers were the biggest group of representatives of one language, but they (alongside with the Serbo-Croatian speakers) for example also use the Latin script in more places than representatives of other languages and they generally use the Latin script more often and in longer sequences of text.

As a kind of a bonus question, I asked the four interviewees why the speakers of their language might have different answers. The Farsi speaking respondent said that in Iran in big cities people would respond differently compared to the countryside. In addition, the age of the respondent could affect the answers. The Russian speaking interviewee mentioned that country of residence could affect the answers. Russian speakers living in a country where the Cyrillic script is widely used might respond differently than those Russian speakers who live in countries where the Cyrillic keyboards are not easily available. The Cantonese speaking interviewee said that younger people in Hong Kong would answer differently than her, since young people nowadays like to use Cantonese more exclusively than older people. According to the Thai speaking interviewee the usage of the Latin script depends on life experience and knowledge of the Thai alphabet. Those who cannot write Thai in Thai alphabet need to write some words in the Latin alphabet. The variety of explanations by the interviewees to the question reflects well the fact that there are dozens of reasons for people to choose to write in the Latin script.

Since the majority of all the questionnaire respondents had started to use the Latin script during primary school years, it is possible that this is also the time when many of the people have had the first contact with the Latin script, for example, when learning English.

In the study by Lee (2007: 188) on Hong Kongese Cantonese speakers, 68% of the respondents preferred to mix Chinese and English when communicating online while 30% used only English and none preferred to use only Chinese. The data of this research does not directly indicate the same results, but it can be agreed that mixing languages and writing systems is fairly popular in online writing. This study shows that abbreviations and loanwords from English written in the Latin alphabet are a common phenomenon in the writing of many of the respondents.



## 7 Conclusions

In this chapter, I answer to the research questions, go through the hypotheses and present the conclusions that can be made from this research. I also introduce some ideas for future studies related to the topic of this study.

The research questions of this research were 1) Do the writers use the Latin script if their language's official writing system is a non-Latin script? 2) If yes, when and why do they do that? The conclusion is that over a half of the respondents use the Latin alphabet and the language they write has little effect on this. Out of the different languages, the Serbian and Serbo-Croatian speakers use the Latin script the most, which is the consequence of the digraphic situation in Serbia. Other active language communities are Bulgarian speakers, Arabic (including all variants) speakers and Mandarin Chinese speakers in mainland China. Those individuals who do use the Latin script, use it because of laziness to switch keyboards, the desire to write quickly and informally. The Latin script is also needed in teaching situations, in international communication and when travelling abroad.

The first research hypothesis was that all the respondents either sometimes write their language in the Latin alphabet or they have seen it written in the Latin alphabet somewhere by others. The second hypothesis was that the main domain of the Latin script use is writing online or with an electronic device and chatting with friends. The third hypothesis was that in some cases writing with the Latin script is voluntary or even desirable but, in some cases the Latin script is used only because the proper keyboard is not available.

The first hypothesis did not actualise because nine respondents, who are speakers of Mandarin Chinese, Cantonese Chinese and Marathi, say that they do not use the Latin script themselves, nor have they seen it used for their language by others. Two respondents, one Mandarin and one Cantonese speaker, use the Latin script themselves, but have not seen the Latin script used by others. All the rest of respondents have at least seen the Latin script used by others. Even though the first hypothesis did not actualise, the second and the third hypothesis turned out to be correct.

From the quantitative data it can be concluded that if a person writes her/his language in the Latin script, there is a high chance that s/he has seen it used by others too (as seen in Table 4 Chapter 4.1.2.). The least likely scenario is that a person writes her/his language in the Latin

script but has not seen it used by others. A clear majority of the respondents like to write their language in the official non-Latin writing system, while only around a half of the respondents like to write their language in the Latin script. Therefore it can be concluded that most people prefer to write in the non-Latin script even though it might sometimes feel difficult or slow to use.

The respondents of “Cyrillic” group use the Latin script quite a lot. The Serbian and Serbo-Croatian speakers use the Latin script in all kinds of domains, while the other languages of the group are more selective with where they use the Latin script. Another clear contrast between the Serbian or Serbo-Croatian speakers and the speakers of other languages in the group “Cyrillic” is that the Serbian and Serbo-Croatian speakers use the Latin script a lot more frequently than others.

In the “Arabic” group the answers of the respondents are fairly consistent even though the group represented eight different languages and language variants. However, when it comes to the perception on the use of the Latin script, the opinions vary a lot. The Arabic script writers use the Latin script mostly if the right script is not available.

In the “Chinese” group, the respondents have various answers in several questions, which makes this group of respondents quite diverse. There are several ways to write the Chinese characters, both simplified and traditional, which reflects in the answers of the three major Chinese language communities. The Latin letters are mostly used in typing Pinyin or when the writer does not remember how to write the correct character.

The answers of the Thai script users are in some parts very unanimous and in some parts different from each other. The Thai speakers do not generally feel the need to use the Latin script, since Thai is considered more suitable for writing Thai. The Latin letters are used in Thai as “Karaoke language”, often only to give pronunciation hints to people who do not know the Thai letters enough.

The group of “Others” is the most diverse and incoherent group which resulted into many different answers. However, in the cases of languages that are represented by more than one speaker, the respondents are fairly unanimous with their answers. Coincidence may play a role in this, but the similar replies give a hint of what a greater representation of those languages might tell.

The study was executed by generalising the writing systems and dividing them into five groups but the results turned out to be very inconsistent within those groups. The general finding of this research is that many people choose to use the Latin script out of individual reasons. Many respondents live abroad and are fluent in the English language, which emphasise the international and highly educated traits that contribute to the results. The individuality of answers is visible within the five script groups as well as within the speakers of one language.

Even though the data is quite dispersed between languages, countries and linguistic backgrounds, there are some tendencies in the use of the Latin script. One of the reasons to use the Latin script is that the writer is used to it or s/he has a habit to write with it. Familiarity of a certain keyboard is often the reason for choosing to use it. For some people typing with the Latin script is easier or faster than with the official writing system. Also, understandability of the text is a factor that affects the choice of a script. In some languages, the Latin script can be used for other people to understand (for example Serbia with neighbour countries as mentioned in Chapter 5.1.2) and in some languages the Latin alphabet is can be used as a secret code (for example United Arab Emirates teenagers as mentioned in Chapter 5.2).

The Latin alphabet has a lot of different functions. Sometimes it is used in formal settings and sometimes in informal settings. For some language communities and individuals, the Latin alphabet is a neutral writing system and for some language communities it evokes positive or negative emotions and sentiments. Sometimes the Latin script is the final option if no other keyboard is available and some situations it is the preferred writing system.

According to the data gathered for this study it seems that the country of residence does not affect whether the respondent uses the Latin script or not when writing their language. People speaking the same language and living in the same country had answered both “yes” and “no” to the question about their possible Latin script use. Both answers were distributed unpredictably also in the cases where people speak the same language but live in different countries. In previous literature (Ivković 2015) it is stated that people living in diaspora need to adjust their writing to the keyboards available which can result to the use of the Latin script. However, the results derived from the data might indicate that people living in diaspora have the necessary keyboards available and that in some cases the keyboard unavailability has not been the reason to use the Latin script.

The phonology of languages and the suitability of writing systems is not discussed in this study. However, it is possible, that the phonology of a language affects the speakers' perception on whether the language is suitable to be written in the Latin script. For example, tonal languages can be tricky to write in the Latin script since often the tones should be separately marked by diacritics, numbers or other conventions. This fact might have also affected the answers of the respondents.

Fundamentally, people choose the script according to the aims of the text and the writing situation. Some mention that they choose the Latin script because they are expected to do so in certain situations. On the other hand, as indicated also by Tagg & Seargeant (2012: 210–212) and (Silva 2013: 147–156) the switching of a script can work as a slightly rebellious act: on the Internet, people feel they are allowed to be playful and the Latin alphabet is one tool to express the playfulness and the freedom to write as they will.

This study inspired some ideas for future research. The suitability of keyboards on different devices for writing different languages should be studied more since there are still some things that should be improved (for example Farsi word prediction on iPhone seems to be non-existent). It would be also relevant to study what kind of consequences the script unavailability has to minority languages. As stated by Anderson (2005: 28) people might even stop writing their language completely if the wanted script is not easily available for them.

In another study similar to this one, it would be good idea to use written text samples as data. This way the research would be based on actual examples of script-switching instead of self-evaluation of the respondent. This kind of study on the use of writing systems could also be made about just one language group, which would give more detailed information about the use of a certain writing system in a specific language group.

A possible question (and also a future research tip) for the respondents who do not write their language in the Latin script, would be “why not?” This research concentrates on the “yes”-side of things and finds out why to choose the Latin script but it would be also interesting to see the other side of the coin.

## References

- Anderson, Deborah. 2005. Global linguistic diversity for the Internet. *Communications of the ACM*. Volume 48, Issue 1, pp. 27–28.
- Androutsopoulos, Jannis. 2012. Chapter 14. “Greeklish”: Transliteration practice and discourse in the context of computer-mediated digraphia. In Jaffe, Alexandra M.; Androutsopoulos, Jannis; Johnson, Sally & Sebba, Mark (eds.) *Orthography as social action: scripts, spelling, identity and power*. pp. 359–392. De Gruyter, Inc.
- Angermeyer, Philipp. 2005. Spelling bilingualism: Script choice in Russian American classified ads and signage. In *Language in Society* 2005 Sept. Volume 34, Issue 4, pp.493–531.
- ASCII. 2019: *Engineering and Technology History Wiki*. <<https://ethw.org/ASCII>>. Last accessed 2019-2-24.
- Axelsson, Ann-Sofie ; Åsa Abelin & Ralph Schroeder. 2007. Anyone Speak Swedish? Tolerance for Language Shifting in Graphical Multiuser Virtual Environments. In Danet, Brenda & Herring, Susan C. (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. pp. 362–381. Oxford: Oxford University Press.
- Backhaus, Peter. 2006: *Linguistic Landscapes: A Comparative Study of Urban Multilingualism in Tokyo*. Channel View Publications.
- Celine. 2017. "Difference Between Urdu and Hindi." *DifferenceBetween.net*. September 28, 2017 <<http://www.differencebetween.net/miscellaneous/culture-miscellaneous/difference-between-urdu-and-hindi/>>. Last accessed 2019-3-30.
- Choksi, Nishaant. 2015. Surface politics: Scaling multiscriptality in an Indian Village market. In Das, Sonia & LaDousa, Chaise (eds.) 2015: *Journal of Linguistic Anthropology*. Volume 25, Issue 1, pp. 1–24.
- Connection Software. 2018. *Character sets for sending bulk SMS text messages*. <<https://www.csoft.co.uk/support/character-sets>> Last accessed 2019-2-24.
- Coulmas, Florian. 2003. *Writing Systems – An introduction to their linguistic analysis*. pp. 231–234. Cambridge: Cambridge University Press.
- Danet, Brenda & Herring, Susan C. 2007. Introduction – Welcome to the Multilingual Internet. In Danet, Brenda & Herring, Susan C. (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. pp. 3–39 Oxford: Oxford University Press.
- DeFrancis, John. 1984. Digraphia. In *Word*. Volume 35, Issue 1, pp.59–66.
- Eisenberg, Leora. 2018. Almaty youth on the future of the Kazakh alphabet. In *Voices on Central Asia*. <<http://voicesoncentralasia.org/almaty-youth-on-the-future-of-the-kazakh-alphabet/>> Last accessed 2019-2-24

- Ellis, Nick C. et al. 2004. The effects of orthographic depth on learning to read alphabetic, syllabic, and logographic scripts. In *Reading Research Quarterly*. Volume 39. Number 4. International Literacy Association.
- Grivelet, Stéphane. 2001a. Introduction. In Duchêne, Alexandre & Coulmas, Florian (eds.). *International Journal of the Sociology of Language*. Volume 2001, Issue 150. pp. 1–10.
- Grivelet, Stéphane 2001b. Digraphia in Mongolia. In Duchêne, Alexandre & Coulmas, Florian (eds.) *International Journal of the Sociology of Language*. Volume 2001, Issue 150. pp. 75–93
- Hansell, Mark. 2015. Roman Letters in Chinese Writing. In *Encyclopedia of Chinese Language and Linguistics*. Brill Online Reference Works.  
<[https://referenceworks.brillonline.com/entries/encyclopedia-of-chinese-language-and-linguistics/roman-letters-in-chinese-writing-COM\\_00000366](https://referenceworks.brillonline.com/entries/encyclopedia-of-chinese-language-and-linguistics/roman-letters-in-chinese-writing-COM_00000366)> Last accessed 2019-3-30.
- Hämeen-Anttila, Paula. 2016. *Elliniká h ellhvika? Latinalaisilla aakkosilla kirjoitettu kreikka internet-keskusteluissa*. MA thesis. University of Helsinki.
- Herring, S. C. 2007. A faceted classification scheme for computer-mediated discourse. In *Language@Internet*. Volume 4, Article 1.  
<<http://www.languageatinternet.org/articles/2007/761>> Last accessed 2019-2-24.
- Hirsjärvi, Sirkka & Hurme, Helena. 2000. *Tutkimushaastattelu: teemahaastattelun teoria ja käytäntö*. Helsinki: Yliopistopaino.
- Hosken, Martin. 2003. Creating an Orthography Description. *SIL Literacy and Education*.  
<<https://www.sil.org/literacy-and-education/resources-developing-orthographies>> Last accessed 2019-2-24.
- ISO (International Organization for Standardization) n.d. *Standards catalogue*. <<https://www.iso.org/ics/01.140.10/x/>> Last accessed 2019-2-24.
- Ivković, Dejan. 2013. Pragmatics meets ideology: Digraphia and non-standard orthographic practices in Serbian online news forums. *Journal of Language and Politics*. Volume 12, Issue 3, pp.335–356.
- Ivković, Dejan. 2015. CyberLatinica: A Comparative Analysis of Latinization in Internet Slavic. In *Language@Internet*. Volume 12, Article 2.  
<<http://www.languageatinternet.org/articles/2015/ivkovic2>> Last accessed 2019-2-24.
- Koutsogiannis, Dimitris. 2007. Greeklish and Greekness: Trends and Discourses of “Glocalness”. In Brenda Danet & Herring, Susan C. (edit). *The Multilingual Internet: Language, Culture, and Communication Online*. pp. 142–158. Oxford: Oxford University Press.
- KvantiMOTV. <<https://www.fsd.uta.fi/menetelmaopetus/kyselylomake/laatiminen.html>> Last accessed 2018-10-29.

Lee, Carmen K. M. 2007. Linguistic Features of Email and ICQ Instant Messaging in Hong Kong. In Danet, Brenda & Herring, Susan C. (eds.) *The Multilingual Internet: Language, Culture, and Communication Online*. pp. 185–205.

Maxfield, Max. 2011. How it was: ASCII, EBCDIC, ISO, and Unicode. In *EETimes* 10.07.2011. <[https://www.eetimes.com/author.asp?section\\_id=216&doc\\_id=1285465](https://www.eetimes.com/author.asp?section_id=216&doc_id=1285465)> Last accessed 2019-2-24.

Palfreyman, David & Al Khalil, Muhamed. 2007. “A Funky Language for Teenzz to Use” Representing Gulf Arabic in Instant Messaging. In Danet, Brenda & Herring, Susan C. (eds.) *The Multilingual Internet: Language, Culture, and Communication Online*. pp. 43–62. Oxford: Oxford University Press.

Panović, Ivan. 2018. ‘You don't have enough letters to make this noise’: Arabic speakers' creative engagements with the Roman script. In Deumert, Ana & Swann, Joan (eds.): *The Sociolinguistics of Everyday Linguistic Creativity*. Volume 65, January 2018, pp. 70–81.

Sampson, Geoffrey. 2015. *Writing systems*. Equinox. Second edition.

Qiu, Phil. 2008. A Forbidden City of words. In *The Guardian*. 2008-7-28. <<https://www.news.com.au/technology/online/censorship/chinas-war-on-words-anything-be-it-a-phrase-or-picture-that-can-be-used-to-insult-xi-has-been-banned/news-story/a8e5a9d558b3ed0465e1fc020e2d6c2c>> Last accessed 2019-3-12.

Silva, Cláudia. 2013. Writing in Portuguese chats :) – A new wrtng systm? In Susanne R. Borgwaldt & Terry Joyce (eds.), *Typology of Writing Systems*, pp. 147–159. Amsterdam: John Benjamins Publishing Co.

Singh, Udaya Narayana. 2001. Multiscriptality in South Asia and language development. In Duchêne, Alexandre & Coulmas, Florian (eds.): *International Journal of the Sociology of Language*. Volume 2001, Issue 150, pp. 61–74.

Spasov, Orlin. 2012. Contesting Bulgaria's past through new media: Latin, Cyrillic and Politics. In *Europe-Asia studies*. Volume 64, Number 8, pp. 1486–1504. Routledge.

Su, Peicheng. 2001. Digraphia: a strategy for Chinese characters for the twenty-first century. In *International Journal of the Sociology of Language 2001* (Aug). Volume 150, pp. 109–124

Su, His-Yao. 2007: The Multilingual and Multiorthographic Taiwan-Based Internet. Creative Uses of Writing Systems on College-Affiliated BBSs. In Brenda Danet & Susan C. Herring (eds.): *The Multilingual Internet: Language, Culture, and Communication Online*. pp. 66–86. Oxford: Oxford University Press.

Tagg, Caroline & Seargeant, Philip. 2012. Writing systems at play in Thai-English online interactions. In Bassetti, Bene & Vaid, Jyotsna (eds.): *Writing systems Research*. Volume 4, Issue 2, pp. 195–213.

Tseliga, Theodora. 2007. “It's All Greeklsh to Me!” In Brenda Danet & Susan C. Herring (toim.), *The Multilingual Internet: Language, Culture, and Communication Online*. pp. 116–141. Oxford: Oxford University Press.

UNGEGN. 2014. *Report on the Current Status of United Nations Romanization Systems for Geographical Names*. <<http://www.eki.ee/wgrs/>> Last accessed on 2019-2-24.

Unicode Standard. 2018. *Supported Scripts*. <<https://www.unicode.org/standard/supported.html>> Last accessed on 2019-2-24.

Usanova, Irina. 2019. *Biscriptuality: Writing Skills among German-Russian Adolescents*. John Benjamins Publishing Company.

World Text n.d. *How long is an SMS text message?* <<https://www.world-text.com/docs/sms-length.php>> Last accessed 2019-2-24.

Zentgraf, David C. 2015. *What every programmer absolutely, positively needs to know about encodings and character sets to work with text*. <<http://kunststube.net/encoding/>> Last accessed 2019-2-24.



## Attachment 1

This is a copy of the online questionnaire to which the respondents gave their answers. All the quantitative data and some of the qualitative data was gathered with this questionnaire. The questions marked with an asterisk (\*) were obligatory. The informants filled in this form in their own free time and without compensation.

### Research questionnaire

#### Typological research on the use of writing systems

This is a research questionnaire for a master's thesis in general linguistics (Department of Languages, University of Helsinki, Finland). The topic of this research is the use of different writing systems all over the world.

In order to participate in this research you should be a native speaker (or otherwise proficient speaker) of a language that uses a non-Latin alphabet based writing system. Languages like English, Spanish and French utilise the Latin alphabet writing system. The non-Latin writing systems are used by languages like Arabic, Mandarin Chinese and Thai.

The questionnaire is anonymous and the IP address of your computer will not be saved. The data collected with this questionnaire will be used for research purposes only. Because of research ethical reasons you should be of legal age in your country of residence. Please answer all the questions you can and as thoroughly as possible.

Thank you for the participation! Feel free to share the link to the questionnaire with your friends. You can contact the researcher Anni Luoma by e-mail: [anni.luoma@helsinki.fi](mailto:anni.luoma@helsinki.fi).

\* 1. Age

--Select--

\* 2. Gender

- ☐ man  
☐ woman  
☐ other

\* 3. Education

- ☐ compulsory school  
☐ college/high school/vocational school  
☐ university student  
☐ university graduate  
☐ other, please specify:

\* 4a. Occupation

--Select--

4b. If you chose "other" on the list of the question 4a, please specify your occupation here.

\* 5a. Nationality

--Select--

5b. Here you can comment on your answer to 5a, if you wish. For example, if you chose "other", you may specify your answer here.

\* 6. Country of residence

--Select--

## \*7a. Language

--Select--

7b. Here you can comment on your choice. If your language was not listed in 7a, you can write it here.

## \*8a. Writing system of your language

--Select--

8b. Here you can comment on your choice. If the writing system of your language was not listed in 8a, you can write it here.

- \*9. Is the Latin script one of the official writing systems along with a non-Latin writing system in your language?
- ☐ yes  
☐ no  
☐ I don't know

- \*10a. Do you ever use the Latin alphabet when writing your language?
- ☐ yes  
☐ no

*If you answered "no" to question 10a, please proceed to question 20a.*

- 10b. If you chose "yes", how often do you use the Latin script to write your language?
- ☐ almost never  
☐ at least once a month  
☐ at least once a week  
☐ every day

11. When did you start using the Latin alphabet when writing your own language?
- ☐ 4-7 years old  
☐ 8-11 years old  
☐ 12-15 years old  
☐ 16-19 years old  
☐ 20+ years old  
☐ 30+ years old  
☐ 40+ years old  
☐ 50+ years old  
☐ 60+ years old  
☐ 70+ years old  
☐ 80+ years old  
☐ 90+ years old  
☐ I don't know/remember

## \*12a. Choose the domains where you use the Latin script when writing your language

- ☐ chatting with friends  
☐ chatting with family members  
☐ chatting with strangers on the internet  
☐ posting on social media  
☐ at work assignments  
☐ at school assignments  
☐ writing to foreign friends  
☐ when translating my language to some other language  
☐ writing info about myself when visiting abroad (e.g. name and address at hotels)  
☐ I don't know/remember  
☐ other

12b. Here you can tell more about the domains of use of the Latin script. You can write domains that were not given in the 12a question, give some examples or specify your answers.

13a. Choose the reasons why you use the Latin alphabet in these situations (given in 12a).

- ☐ it feels easier than the official writing
- ☐ it is faster than the official writing
- ☐ the binary code of Latin letters is shorter than the one of the official writing
- ☐ other keyboards were not available
- ☐ just for fun
- ☐ others do it too
- ☐ using western alphabet is cool
- ☐ some things are not translated to my language (e.g. CD, Facebook)
- ☐ other

13b. Here you can tell more about the reasons for using the Latin script. You can write reasons that were not given in the 13a question, give some examples or specify your answers.

14a. When you write your language with the Latin alphabet, how much of the text you write with it?

- ☐ just some letters within a word
- ☐ just 1-3 words inside a sentence
- ☐ a complete sentence
- ☐ several sentences
- ☐ the whole text
- ☐ various lengths depending on what is written

14b. Here you can tell more about the length of the texts written in the Latin script or specify your answer otherwise.

15a. Do you ever write your language with the Latin alphabet if you are writing by hand (e.g. shopping list/letters/notes)?

- ☐ yes
- ☐ no

15b. If you chose "yes" to question 15a, please write here some more information about writing by hand in the Latin script. You can write examples of the domains or tell how often you write with the Latin script by hand.

16a. When you use the Latin writing in your language, do you follow some official spelling rules (e.g. standardised romanisation system)?

- ☐ yes
- ☐ no
- ☐ I don't know

16b. Here you can tell more about your answer in 16a and give some examples.

17a. When you use the Latin alphabet in your language, you use the

- ☐ formal standard language
- ☐ spoken /vernacular language
- ☐ both

17b. Here you can tell more about your answer in 17a and give some examples.

18a. Would you use the Latin alphabet when writing official papers in your language? ☐ yes ☐ no

18b. Here you can tell more about your answer in 18a and give some examples.

19a. Do you like writing your language in the Latin alphabet? ☐ yes ☐ no

19b. Here you can tell more about your answer in 19a and give some examples.

20a. Do you like writing your language with the official non-Latin writing system? ☐ yes ☐ no

20b. Here you can tell more about your answer in 20a and give some examples.

21a. Have you seen your language written in the Latin alphabet somewhere? ☐ yes ☐ no

*If you answered "no" to question 21a, please proceed to question 23.*

21b. If you chose "yes", choose the domains where you have seen the Latin alphabet being used in your own language.

- ☐ friends chatting with electronic devices
- ☐ family members chatting with electronic devices
- ☐ strangers on the internet chatting
- ☐ social media posts
- ☐ Latin alphabet is used at work
- ☐ on street signs
- ☐ in TV advertisement
- ☐ in newspaper advertisement
- ☐ in shop products
- ☐ poetry
- ☐ novels
- ☐ religious texts
- ☐ children's books
- ☐ language learning materials
- ☐ hand written texts (e.g. shopping list/letters/notes)
- ☐ other

21c. Here you can tell more about your choices in question 21b and give some examples. If you chose "other", please write here what it would be.

22a. What is usually written in the Latin alphabet in your language by you or by others?

- ☐ words from English
- ☐ words from other foreign languages
- ☐ foreign names of people
- ☐ foreign names of places/buildings/other geographical names
- ☐ pronunciation rules for your language
- ☐ abbreviations
- ☐ anything is possible to write in the Latin alphabet in your language
- ☐ other

22b. Here you can tell more about your choices in question 22a and give some examples. If you chose "other", please write here what it would be.

22c. If you use the Latin alphabet to write loanwords from foreign languages (other than English), what are these languages?

23. What phone brand do you use mostly?

- ☐ Samsung  
☐ Apple/iPhone  
☐ Huawei  
☐ Oppo  
☐ Vivo  
☐ Xiaomi  
☐ LG  
☐ Lenovo  
☐ ZTE  
☐ Alcatel-Lucent  
☐ Nokia/Android  
☐ Nokia/Windows  
☐ HTC

☐ other, please specify:

24a. Do you use a Latin alphabet-based keyboard on electronic devices when writing the script of your language (e.g. pinyin with simplified Chinese)?

- ☐ yes  
☐  
☐ no

24b. If you answered "yes" to 24a, please write here what Latin alphabet based spelling system you use.

25. Do you think the keyboard on your phone is working well when switching between different writing systems?

- ☐ very well  
☐ pretty well  
☐ quite poorly  
☐ very bad  
☐  
☐ varying depending on the situation  
☐ I don't know

26. Do you think the keyboard system for the official writing system of your language is working well in general on all electronic devices?


- ☐ very well  
☐ pretty well  
☐ quite poorly  
☐ very bad  
☐  
☐ varying depending on the situation  
☐ I don't know

27a. Do you think the Latin alphabet keyboard is more easily available than the official alphabet keyboard of your language on phones and other electronic devices?

- ☐ yes  
☐  
☐ no

27b. Here you can tell more about your answer in 27a and give some examples.

28. Is there something else you would like to mention about the use of different writing systems in your language?



☐ \* I accept that my answers will be used for research purposes.

This is the end of the questionnaire. Remember to submit the filled form. Thank you for your answers!

**Proceed**

---

Submit

## Attachment 2

Due to the relatively large size of the data, it was not possible to present all of the datasets in the attachments. Instead of displaying the whole data, here are four arbitrarily selected questionnaire forms filled by the respondents. The examples demonstrate the data format that was analysed for this research.

<b>Number of submission</b>	4	49	54	58
<b>Time of submission</b>	20.12.18 20:18	23.12.18 01:03	24.12.18 00:30	24.12.18 04:50
<b>1. Age</b>	51	33	19	33
<b>2. Gender</b>	woman	man	woman	man
<b>3. Education</b>	other	university graduate	college/high school/vocational school	university graduate
Other, please specify:	PhD			
<b>4a. Occupation</b>	Education/teaching	IT/data communications	Other	Construction
<b>4b. If you chose "other" on the list of the question 4a, please specify your occupation here.</b>			student	
<b>5a. Nationality</b>	Greek	Iranian	Polish	Thai
<b>5b. Here you can comment on your answer to 5a, if you wish. For example, if you chose "other", you may specify your answer here.</b>				
<b>6. Country of residence</b>	Greece	Finland	United Kingdom	Thailand
<b>7a. Language</b>	Greek	Persian/Farsi	Russian	Thai
<b>7b. Here you can comment on your choice. If your language was not listed in 7a, you can write it here.</b>			my native language is polish, however i am fluent in russian and regularly use it - at least once a day.	
<b>8a. Writing system of your language</b>	Greek	Arabic	Cyrillic	Thai

<b>8b. Here you can comment on your choice. If the writing system of your language was not listed in 8a, you can write it here.</b>				
<b>9. Is the Latin script one of the official writing systems along with a non-Latin writing system in your language?</b>	no	no	no	no
<b>10a. Do you ever use the Latin alphabet when writing your language?</b>	no	yes	yes	yes
<b>10b. If you chose "yes", how often do you use the Latin script to write your language?</b>		at least once a week	at least once a month	every day
<b>11. When did you start using the Latin alphabet when writing your own language?</b>		12-15 years old	16-19 years old	4-7 years old
<b>12a. Choose the domains where you use the Latin script when writing your language</b>				
chatting with friends		x		x
chatting with family members				x
chatting with strangers on the internet				x
posting on social media				x
at work assignments				x
at school assignments			x	x
writing to foreign friends		x		
when translating my language to some other language		x		



writing info about myself when visiting abroad (for example name and address at hotels)		x		
I don't know/remember				
other			x	
<b>12b. Here you can tell more about the domains of use of the Latin script. You can write domains that were not given in the 12a question, give some examples or specify your answers.</b>				
<b>13a. Choose the reasons why you use the Latin alphabet in these situations (given in 12a).</b>				
it feels easier than the official writing				x
it is faster than the official writing				x
the binary code of Latin letters is shorter than the one of the official writing				x
other keyboards were not available		x	x	
just for fun				
others do it too				x
using western alphabet is cool				
some things are not translated to my language (for example CD, Facebook)				x
other			x	
<b>13b. Here you can tell more about the reasons for using the Latin script. You can write reasons that were not given in the 13a question, give some examples or specify your answers.</b>			most in my country of residence - england - can't read cyrillic.	

<b>14a. When you write your language with the Latin alphabet, how much of the text you write with it?</b>		just some letters within a word	just 1-3 words inside a sentence	just 1-3 words inside a sentence
<b>14b. Here you can tell more about the length of the texts written in the Latin script or specify your answer otherwise.</b>				
<b>15a. Do you ever write your language with the Latin alphabet if you are writing by hand (for example shopping list/letters/notes)?</b>		no	no	yes
<b>15b. If you chose "yes" to question 15a, please write here some more information about writing by hand in the Latin script. You can write examples of the domains or tell how often you write with the Latin script by hand.</b>				
<b>16a. When you use the Latin writing in your language, do you follow some official spelling rules (for example standardised romanisation system)?</b>		no	no	I don't know
<b>16b. Here you can tell more about your answer in 16a and give some examples.</b>			the way i transcribe cyrillic to latin alphabet is a mixture of both the english and polish orthography	
<b>17a. When you use the Latin alphabet in your language, you use the</b>		spoken/vernacular language	both	spoken/vernacular language

<b>17b. Here you can tell more about your answer in 17a and give some examples.</b>				
<b>18a. Would you use the Latin alphabet when writing official papers in your language?</b>		no	no	no
<b>18b. Here you can tell more about your answer in 18a and give some examples.</b>				
<b>19a. Do you like writing your language in the Latin alphabet?</b>		no	no	
<b>19b. Here you can tell more about your answer in 19a and give some examples.</b>				No preference
<b>20a. Do you like writing your language with the official non-Latin writing system?</b>	yes	yes	yes	
<b>20b. Here you can tell more about your answer in 20a and give some examples.</b>	Historical reasons, ancient alphabet			No preference
<b>21a. Have you seen your language written in the Latin alphabet somewhere?</b>	yes	yes	yes	yes
<b>21b. If you chose "yes", choose the domains where you have seen the Latin alphabet being used in your own language.</b>				
friends chatting with electronic devices	x	x		x
family members chatting with electronic devices	x			x
strangers on the internet chatting	x	x	x	x
social media posts	x	x		x

Latin alphabet is used at work				
on street signs	x	x	x	x
in TV advertisement	x			x
in newspaper advertisement	x	x		x
in shop products	x	x		x
poetry			x	
novels			x	
religious texts				
children's books				x
language learning materials		x	x	x
hand written texts (for example shopping list/letters/notes)				
other				
<b>21c. Here you can tell more about your choices in question 21b and give some examples. If you chose "other", please write here what it would be.</b>				
<b>22a. What is usually written in the Latin alphabet in your language by you or by others?</b>				
words from English	x			x
words from other foreign languages	x			
foreign names of people	x		x	x
foreign names of places/buildings/other geographical names	x		x	x
pronunciation rules for your language	x			
abbreviations	x			
anything is possible to write in the Latin alphabet in your language		x		
other				
<b>22b. Here you can tell more about your choices in question 22a and give some examples. If you chose "other", please write here what it would be.</b>				

<b>22c. If you use the Latin alphabet to write loanwords from foreign languages (other than English), what are these languages?</b>	french, german, italian, spanish,danish			
<b>23. What phone brand do you use mostly?</b>	Other	Apple	Apple	Apple
Other, please specify:	Blackberry			
<b>24a. Do you use a Latin alphabet-based keyboard on electronic devices when writing the script of your language (for example pinyin with simplified Chinese)?</b>	no	no	yes	no
<b>24b. If you answered "yes" to 24a, please write here what Latin alphabet based spelling system you use.</b>			English (US)	
<b>25. Do you think the keyboard on your phone is working well when switching between different writing systems?</b>	very well	quite poorly	very well	pretty well
<b>26. Do you think the keyboard system for the official writing system of your language is working well in general on all electronic devices?</b>	pretty well	quite poorly	pretty well	pretty well
<b>27a. Do you think the Latin alphabet keyboard is more easily available than the official alphabet keyboard of your language on phones and other electronic devices?</b>	yes	yes	yes	yes

<b>27b. Here you can tell more about your answer in 27a and give some examples.</b>			if i were to use a non-official keyboard, it would mostly likely only allow for a latin english keyboard	
<b>28. Is there something else you would like to mention about the use of different writing systems in your language?</b>				
<b>I accept that my answers will be used for research purposes.</b>	x	x	x	x